

CHEMICAL REVIEWS

GERALD WENDT, EDITOR-IN-CHIEF

*Dean of the School of Chemistry and Physics of the Pennsylvania
State College, State College, Pennsylvania*

VOLUME V

BALTIMORE
1928

CONTENTS

NUMBER 1, FEBRUARY, 1928

Active Glucose. P. A. LEVENE.	1
The Mineralogy of Hafnium. O. IVAN LEE	17
The Radiation Hypothesis of Chemical Reaction. FARRINGTON DANIELS.	39
Catalytic Actions. ERIC K. RIDEAL	67

NUMBER 2, JUNE, 1928

The Structure of Atoms as a Periodic Property and Its Relation to Valence and Ion-Formation. ARTHUR A. NOYES AND ARNOLD O. BECKMAN.	85
Line Spectra and the Periodic Arrangement of the Elements. SAUL DUSHMAN	109
The Application of the Quantum Mechanics to the Structure of the Hydrogen Molecule and Hydrogen Molecule-Ion and to Related Problems. LINUS PAULING.	173
The Helix Chemical. B. K. EMERSON.	215

NUMBER 3, SEPTEMBER, 1928

Acid and Basic Catalysis. J. N. BRÖNSTED	231
Free Inorganic Radicals. P. WALDEN AND L. F. AUDRIETH	339

NUMBER 4, DECEMBER, 1928

A Symposium on Atomic Structure and Valence. An Introduction. GEORGE L. CLARK.	361
Recent Progress in the Study of the Structure of the Nucleus. C. LIND	365
The Synthesis of Atoms, the Whole Number Rule, and the Periodic System of the Atomic Species. WILLIAM D. HARKINS	371
Atomic Structure as Revealed by Crystal Reflection of X-Rays. G. E. M. JAUNCEY.	437
Dissociation of Molecules as Disclosed by Band-Spectra. KARL K. DARROW.	451

The New Quantum Mechanics. J. H. VAN VLECK	467 [*]
The Electron Theory of Valence. WORTH H. RODEBUSH	509
Valence Forces as Calculated from Specific Heats and Absorption Spectra. DONALD H. ANDREWS	533
The Relation of Shared Electrons to Potential and Absolute Polar Valences. WILLIAM ALBERT NOYES	549
A Pragmatic System of Notation of Electronic Valence Conceptions in Chemical Formulas. HARRY SHIPLEY FRY	557
The Theory of Partial Polarity of the Ethylene Bond and the Existence of Electro-Isomerism. M. S. KHARASCH AND F. R. DARKIS	571
The Broader Aspects of Valence: Its Application to Colloid Compounds. VICTOR COFMAN	603

ACTIVE GLUCOSE*

P. A. LEVENE

Laboratories of The Rockefeller Institute for Medical Research, New York

Glucose is one of the most important substances in the economy of the world. In the living organism it is the source of animal heat. In every-day life it is important as the material from which alcohol, acetone, lactic and tartaric acids are made; furthermore, it is the substance from which are built the molecules of important products such as starches, gums, cellulose.

It is easy then to understand the reasons behind the search for an explanation of the mechanism by which glucose is transformed into these important substances. As a result of the research directed toward this goal, there have been advanced novel theories concerning the forces which serve to combine the smaller molecules into the larger and concerning those forces which lead to the disintegration of the small molecules into still smaller. One of the new conceptions is that of "active" glucose. A concrete, final, undisputed theory of the nature or of the structure of this form of glucose has not yet been furnished. The concept arose as a matter of expediency in order to explain some puzzling observations on the conduct not only of glucose but also of a much simpler sugar, namely, of a triose, a substance which can be derived from glucose and which may condense into glucose.

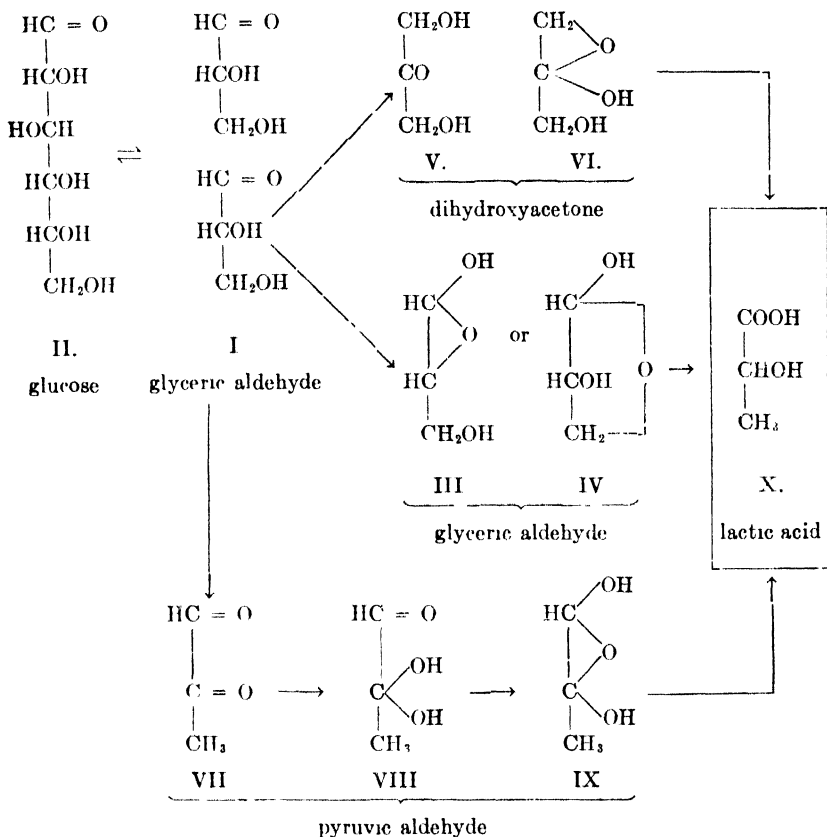
This is a very fortunate occurrence for the reason that the ideas concerning the "active" form of triose are easier to test experimentally than those regarding the "active" form of glucose.

The gradual evolution of the idea regarding triose, therefore, will be presented first by way of introduction to the more involved topic of "active" glucose.

* Presented at the seventy-fourth meeting of the American Chemical Society, Detroit, Michigan, September 6, 1927.

ACTIVE TRIOSE

From the following formulae



it is readily seen that glyceric aldehyde may, on one hand, condense into glucose and, on the other, through intramolecular rearrangements, assume a variety of forms of which pyruvic aldehyde and lactic acid are of special interest.

Theoretically, both pyruvic aldehyde and lactic acid can readily be transformed into acetaldehyde and a substance readily convertible into carbon dioxide. Further, acetaldehyde can easily be either reduced to alcohol or oxidized to carbon dioxide and water. Thus, the chemical theory permits us to outline the

individual phases or steps through which a molecule of glucose has to pass on its way to carbon dioxide and water or on its way to carbon dioxide and alcohol and permits us to formulate a satisfactory theory of fermentation and of the biological oxidation of sugar. Indeed, Nef and Wohl have done so (1).

The experimental evidence in support of the theory had a slow development. The formation of lactic acid from glucose was observed by Reichard in 1870, more recently by Kiliani, still more recently by Meisenheimer, and many times since (2).

The discovery of pyruvic aldehyde among the products of dissociation of glucose in alkaline aqueous solution is of more recent date. It was made first by Pinkus in 1898, later by W. Löb, by Dakin and Janney, by Neuberg and Oertel and by others (3). Finally, the conversion of pyruvic aldehyde into lactic acid through the action of alkali has been demonstrated by Evans and his coworkers (4).

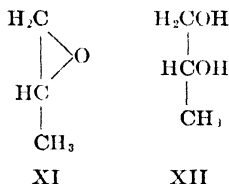
Gradually evidence accumulated to show that biological agents were capable of transforming glucose into lactic acid and to show that its precursor is pyruvic aldehyde. Many investigators, on the basis of animal experimentation, had accepted the view that lactic acid is an intermediary substance in the process of biological utilization of sugar (5). By a more direct and rigorous method the biological conversion of glucose into lactic acid was shown simultaneously and independently by Slosse, by Levene and Meyer, and by Embden and coworkers (6). Levene and Meyer have shown also on experimental grounds that glucose in its conversion into lactic acid passes through the stage of pyruvic aldehyde (7). Dakin and Dudley, Neuberg, and Levene and Meyer have shown independently and practically simultaneously that methyl glyoxal is converted into lactic acid by animal cells or tissues (8). Still later methylglyoxal was isolated from among the products of bacterial action.

Thus it seemed that the mechanism of alcoholic fermentation outlined on purely theoretical grounds had been well supported by experimental evidence. There remained to test the fermentability of pyruvic aldehyde and of lactic acid. It was expected that, as intermediary products, they should ferment at a higher

rate than glucose. The test of the fermentability of pyruvic aldehyde and of lactic acid and the test of their biological oxidation led to very disappointing results. It was found that not one of the several possible forms of pyruvic aldehyde or the substances formed from it through intramolecular rearrangement fermented with the same velocity as glucose (9). In order, therefore, to reconcile the older theory with the newer facts, an assumption was made of the transitory existence of an "active" form of the triose. A concrete formulation, however, of the structure of the "active" form had at that time not been advanced.

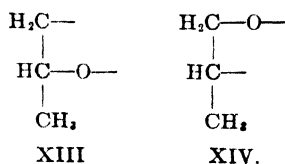
The key for further development of the idea into a more concrete formula was found in the chemical theory.

Scrutinizing all the isomeric forms of a triose, our attention centered on the cyclic structures. The existence of the cyclic forms was demonstrated by the preparation of the corresponding glucosides by Bergmann (10). The free cyclic trioses have not yet been prepared. However, analogous cyclic structures are known for certain organic oxides. Of these, one, propylene oxide, has recently been studied by Levene and Walti (11).

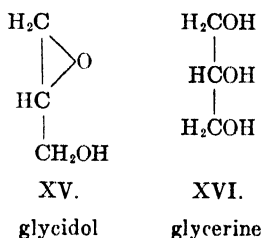


Propylene oxide (XI) is characterized by much greater reactivity than propylene glycol (XII) from which it is readily formed and into which it is transformed still more readily.

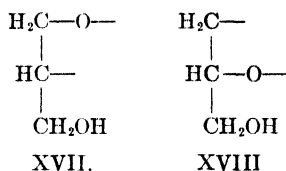
Propylene oxide is easily converted into propionic aldehyde which is formed by an intramolecular rearrangement analogous to that observed in trioses and expressed in formulae from (I) to (X). It had been assumed by Nef that in solution propylene oxide may exist in one of the following two forms.



More definite experimental evidence in favor of the transitory existence of such forms is furnished by recent work of Levene and Walti (11). It has been shown also that oxides have a great tendency towards condensation in a manner similar to that by which simple sugars condense to di- and polysaccharides (12). The conduct of the cyclic substance glycidol and of glycerine is very similar to that of propylene oxide and propylene glycol.



Also, of these two substances, the first is the more reactive. Glycidol, upon opening of the ring, is capable of a transitory existence in the unsaturated forms (XVII) or (XVIII).



These unsaturated forms, as free radicles, possess much greater molecular asymmetry and because of it much greater reactivity than the saturated forms and to them one may justly apply the term "active." Thus, each cyclic triose may exist in its own "active" form, namely, in that which has a transitory existence upon the opening of the ring.

"ACTIVE" GLUCOSE

The evolution of the concept of "active" glucose passed through the same phases as that of "active" triose.

The biological observations which led to the belief in the existence of an "active" form of glucose are the following.

First. The different sugars ferment at different velocities and yet the intermediate and final products of their fermentations are identical. Thus, it was found that when a mixture of fermentable hexoses is exposed to the action of yeast, a marked difference is observed in the rate of fermentation of individual isomers (13).

It is possible that in this respect individual yeasts show different behavior for the reason that the results reported by individual workers are not always in agreement with each other. Thus, according to Fernbach (14), all yeasts ferment fructose with greater velocity in the early phases of fermentation, but the velocity lags behind that of glucose in the later stages of fermentation. However, by a suitable adjustment of the hydroxyl ion concentration, the more rapid fermentability of fructose may be maintained until the very end. If a mixture of glucose and fructose is employed, then according to Willstätter, the rate of fermentation of fructose to glucose is 3:1 (15).

If instead of living yeast, zymon or yeast juice is employed, then, according to Harden (16), fructose has the highest rate of fermentation in the early stages, the second place belongs to mannose, but in the later stages the rate of fermentation of mannose falls behind that of glucose.

Also animal cells and tissues seem to discriminate between the individual hexoses. Thus, Levene and Meyer have observed that the yield of lactic acid from mannose is lower than that from glucose and fructose when living leucocytes or aseptic kidney tissue (17) is allowed to act on them, and according to Laquer (18) the rates of glycolysis in muscle tissue of the three hexoses—fructose, mannose and glucose—are the following. 0.484, 0.395, 0.327. According to Warburg, the cells of various normal and pathological tissues oxidize the three hexoses in the following order, glucose, mannose and fructose, the ratios being 24, 22, 3 (19). Galactose is the least fermentable hexose and in fact, yeast has to undergo special training in order to ferment this sugar.

All the above quoted observations are interpreted to mean that the various sugars are transformed into a form common to all

which then undergoes further decomposition, and that this form is the so-called "active" glucose.

Second. Euler and his coworkers have found that glycogen ferments at a higher rate than free glucose (20). More striking is the high rate of dissociation of glycogen by animal tissues. Thus Laquer found that the formation of lactic acid from glycogen proceeds at a higher speed than that from any other sugar. Embden (21) found that after addition of sodium fluoride, glycogen gives a larger yield of lactacidogen than does glucose. Neuberg and Gottschalk found that not only lactic acid, but also the intermediate products of glycolysis, are formed at a higher rate from glycogen than from other sugars (22).

The nearly unanimous conclusion drawn from these observations is that in glycogen glucose is present in a more labile state than in the common forms of glucose.

Third. A disparity was observed between the fall of the reducing power of a fermenting solution and the amount of alcohol and CO_2 formed. The earlier observations in this direction were made by Euler and his coworkers and by Lebedev (23). The interpretation of Euler was that in the early phase of the process an intermediary product was formed which subsequently underwent fermentation. Indeed, Euler sought for ways to depress fermentation without affecting the first phase of the process and by this means to isolate the intermediate substance. His efforts in this direction, however, failed. Harden and Young, Buchner and Meisenheimer have thought to explain the phenomenon by the assumption of reconversion of glucose to a polysaccharide. Their conclusions, however, were contraverted by Euler. Also Lebedev on theoretical grounds disagreed with the theory of synthesis of a polysaccharide. The observations of Lebedev were interesting also for the reason that he observed a greater lowering in the rotatory power of the fermenting solution than that expected on the basis of the amount of sugar converted into alcohol and CO_2 .

The findings of Lebedev bring us close to the observations made on the blood sugar after administration of insulin. In several publications from the most reputed laboratories the claim was made that after administration of insulin the blood sugar acquired

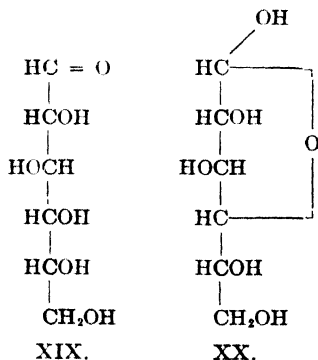
special properties which distinguished it from ordinary glucose, namely, that it had a lower rotatory power than the common forms of glucose and that on standing the rotatory power rose to that of the equilibrium form of common glucose (24).

These claims have been challenged and, to say the least, need corroboration, but whatever the value of the observations may be, they stimulated the belief in the existence of the unstable form of glucose.

Fourth. In the course of fermentation with zymine, several hours pass before the evolution of carbon dioxide begins. It seems as if a period of incubation were required during which the sugar undergoes certain changes necessary for its fermentability. The period of inactivity may be shortened by the addition of some chemical substances, particularly by a substance itself a product of fermentation, namely, a hexosediphosphoric ester. The details of the chemical changes taking place during the incubation period are as yet not known.

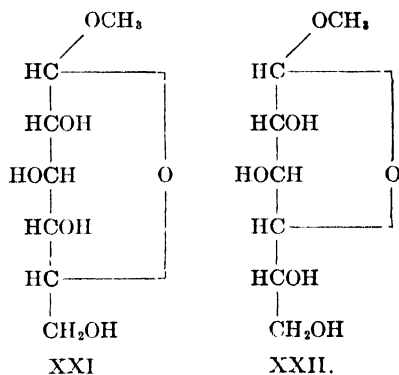
Thus far the problem had been of a purely biological nature. The conclusions had been reached by purely biological methods and lacked concreteness.

For more precise and more accurate conclusions, a deeper insight into the chemistry of sugars is needed. Indeed, the further development of the idea of "active" glucose goes hand in hand with the progress of pure chemistry. In the early days of sugar chemistry it was known that glucose is capable of existence in two forms, one having the structure of an ordinary aldehyde and the other, a cyclic form.



If we had at that period of our knowledge speculated on the nature of an "active" glucose, the choice would have been limited to the two forms only. Indeed, even today, authorities in sugar chemistry are inclined to regard the aldehydic form as the "active" form of glucose.

In recent years, however, heretofore unknown isomers of derivatives of glucose have been discovered. These isomerisms were explained by the assumption of differences in the ring structures. The greatest mass of work was done on the glucosides. Through the work done on these substances by Pryde, Levene and Meyer, Haworth and coworkers and by Levene and Simms, the characters of the different ring structures are now definitely established (25). The common forms of glucosides have the oxygen bridge between the first and the fifth carbon atoms (XXI) and the newly-discovered forms have the oxygen bridge between the first and the fourth carbon atoms (XXII).



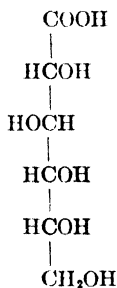
It was proposed to refer to the cyclic forms as lactal forms, naming the first <1,5> lactal and the second <1,4> lactal.

From the facts established for glucosides, it was surmised that free glucose also exists in <1,5> and in <1,4> lactal forms and that the <1,4> lactal form is the more unstable and hence the form to be regarded as the "active" form of glucose.

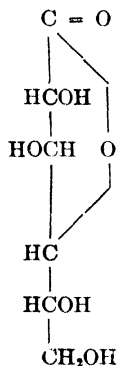
This assumption, plausible as it seemed, at the time when it was announced had little experimental evidence in its support. It is the outstanding problem of the day in sugar chemistry to

unravel the ring structure of the free sugars and to correlate chemical structure with biological conduct.

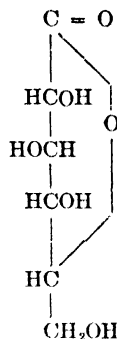
The approach towards the solution of the problem was the following. Sugar acids resemble sugars in their ability to acquire cyclic structures, as shown in the following figures (26).



XXIII



XXIV.



XXV

The advantage of choosing lactones is that, unlike the lactals, they can be prepared easily in crystalline form. It was shown by Levene and Simms that individual lactone forms have different degrees of stability and that peculiarities arising from the ring structure in the case of unsubstituted and in that of substituted acids are practically identical. Thus the conclusion is warranted that also in the case of sugars, there exists an analogy between the properties of the substituted and of the unsubstituted forms.

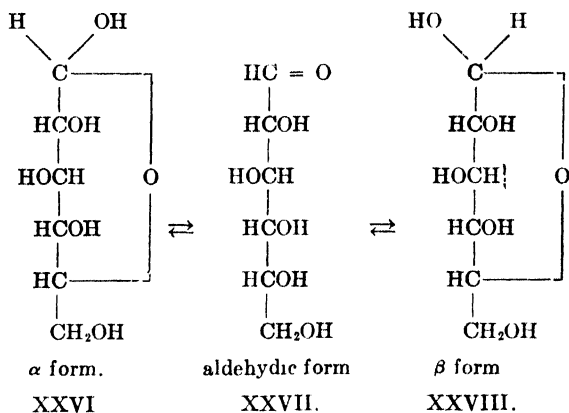
From the observations on substituted and unsubstituted sugar acids the conclusion was formulated that in a solution of a sugar acid, all theoretically possible lactones are formed, but only the more stable are present in measurable quantities and that the proportions of the different lactones are a function of time. In the fresh solutions the unstable forms predominate.

Inasmuch as lactone and lactal formation are analogous processes, the thought was expressed that in a solution of sugar also, all theoretically possible lactals are formed and that at the point of equilibrium the more stable forms predominate. This view is supported by the fact that from the same common glucose,

depending upon the conditions of experiment, three different forms are obtainable. Thus, in methyl alcoholic solution, methyl glucosides are obtained having the $\langle 1,5 \rangle$ and the $\langle 1,4 \rangle$ lactal structures. In acetone solution, only the $\langle 1,4 \rangle$ lactal derivative is formed, whereas in aqueous hydrochloric acid glucose condenses with mercaptans into the noncyclic aldehydic derivatives.

On the basis of these considerations, the assumption of the simultaneous existence in solution of the aldehydic and of several cyclic forms gains in probability.

At the time when this assumption was made by Levene and Simms, there existed no experimental evidence for the view that individual lactals possess different degrees of stability. In sugars and in certain of their derivatives, the stability of a ring may be measured by the phenomenon of mutarotation. In sugars the mutarotation is explained by the change of one isomeric form into another—(α) into (β).



The intermediate is the noncyclic form. Hence the sugars or their derivatives which show a higher velocity of mutarotation are also those in which the opening of the cyclic structures proceeds with higher velocity.

Up to last year there was lacking material of the $\langle 1,5 \rangle$ lactal structure on which the rate of mutarotation could be studied. In course of last year such material was furnished through the

work of Levene and Meyer (27), and today it is quite certain that in free sugars also, the stability of the cyclic structures depends upon the ring structure, and in addition that in free sugars different ring isomers actually exist.

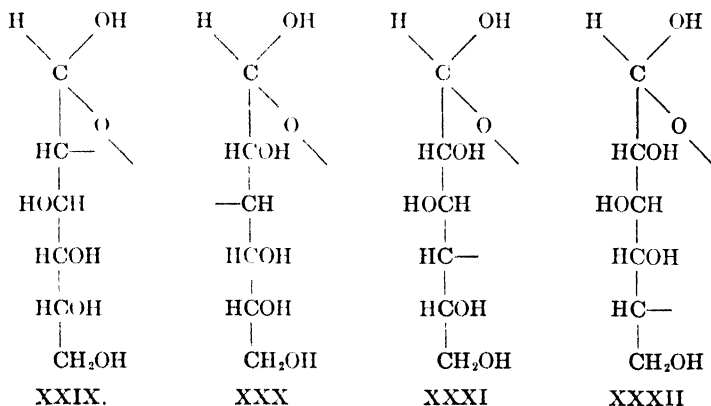
In order to correlate the chemical findings of the stabilities of the different lactal structures with the concept of "active" glucose demanded by the observations on fermentation, it was now necessary to prove that the velocity of fermentation also is influenced by the stability of the lactal structures. Indeed, Willstätter and Sobotka found (28) that α -glucose, (having a more unstable ring structure) ferments with a higher velocity than the β -form.

Thus, it was on one hand shown that in free sugars rings of different stability may exist and on the other, that the rate of fermentation is influenced by the stability of the ring.

Here again we see that after certain facts have been established by the chemical method and after certain conclusions have been drawn from them, it is pertinent to test the chemical theories by the biological method.

If we now return to the view that the "active" form is the one possessing the properties of a free radicle, it will at once be evident that the product of fermentation is in a way determined by the character of the ring structure.

A glance at the free radicles derived from some of the theoretically possible lactals



reveals the fact that the radicles (XXX) and (XXXI) are most likely to undergo the cleavage of the six-carbon chain into two three-carbon chains. These forms therefore may be regarded as the "active" forms of glucose in the fermentations forming alcohol, acetone and lactic acid. On the other hand, for fermentations resulting in butyric acid the $<1,2>$ and the $<1,3>$ lactals are the most favored forms. Thus, each form of fermentation has its own "active" form of glucose.

Granting that the views here expressed are correct, one question still remains unanswered, namely, as to the nature of the mechanism by which in a solution of glucose the concentration of the favored form increases above the normal concentration. Normally, in a fresh aqueous solution of glucose, all lactal forms coexist and after some interval of time, gradually the most stable form predominates.

How then is the concentration of the less stable form increased over that of the more stable forms? The simplest mechanism would be one by which an enzyme enters in combination with a carbon atom which normally serves for the formation of the oxygen bridge. Thus, if a methyl group is introduced in the hydroxyl of the fifth carbon atom, then the formation of the $<1,5>$ lactals is made impossible. The same would happen if an enzyme had combined in place of the methyl group. Often the nature of the group entering into union with the first carbon atom influences the ring structure of the resulting derivative. Thus, acetone glucose in which the acetone is linked to the first and the second carbon atoms has the $<1,4>$ lactal structure.

Indeed, the assumption of the formation of a complex glucose prior to fermentation acquired an explicit character when it was discovered that in the course of fermentation as well as in the course of the combustion of sugar in the muscle tissue, there takes place a condensation of the glucose with phosphoric acid giving rise to a hexose diphosphate. It was then concluded that in this substance the sugar is contained in a more unstable state. In other words, that either the hexose diphosphate itself or the sugar liberated from it is the unstable form. This contention, however, is not fully in accord with experimental evidence.

Phosphorylated glucoses, whether synthetic or formed on fermentation, are not fermented by living yeast or by zymase at a higher rate than free glucose. However, the failure of the phosphoric ester theory does not exclude the possibility of the formation of another complex substance. It is safer, however, to admit that as yet the mechanism by which the concentration of the unstable lactal structures is increased above the normal is unknown.

Thus, the arguments advanced in this discussion are the following.

1. Observations on the processes of fermentation and of animal combustion of glucose led to the conclusion of the existence of an "active" form of glucose.

2. Observations on organic oxides and glycols bring out the fact that cyclic structures are more active than noncyclic structures, because they are readily transformed for a short but finite interval of time into free radicals.

3. Observations on glucose derivatives show that they are capable of existence in several ring isomers differing from each other in their stability.

4. Observations on sugar acids demonstrate that in aqueous solution all theoretically possible cyclic structures can exist simultaneously.

5. The conclusion is drawn that in aqueous sugar solutions, all theoretically possible cyclic structures coexist in equilibrium with each other.

6. The agents bringing about fermentation produce an increase in concentration of that cyclic form of glucose which gives rise to that radical which is most apt to cause the required dissociation of the glucose molecule, the initial phase in the process being the formation of a free radical.

REFERENCES

- (1) NEF, J. U , *Ann Chem* , **335**, 247 (1904)
WOHL, A , *Biochem Zeit* , **5**, 45 (1907)
- (2) REICHARD, H., *Zeit d Vereins d d Zuckerind* , **29**, 529 (1870).
KILIANI, H , *Ber. deut Chem Ges* , **15**, 136, 701 (1882).
MEISENHEIMER, J , *Ber deut Chem Ges* , **41**, 1009, (1908).

- (3) PINKUS, G , Ber. deut. Chem. Ges , **31**, 31 (1898).
 LÖB, W., Biochem Zeit., **12**, 78, 466 (1908).
 DAKIN, H. D , and JANNEY, N W., J. Biol. Chem , **15**, 177 (1913).
 NEUBERG, C , and OERTEL, W , Biochem. Zeit., **55**, 495 (1913).
- (4) EVANS, W. L , AND Co-WORKERS, J. Am. Chem. Soc., **47**, 3085, 3098, 3102, (1925); **48**, 2665, 2678, 2703 (1926).
- (5) SPIRO, P., Zeit. physiol. Chem., **1**, 111 (1877).
 ARAKI, T., Zeit. physiol. Chem , **15**, 335, 546 (1891); **16**, 201, 453 (1892); **17**, 311 (1893); **19**, 422 (1894).
 ZILLESSEN, H , Zeit. physiol. Chem , **15**, 387 (1891).
 EMBDEN, G., Centralbl f. physiol , **18**, 832 (1905).
 MANDEL, A R , AND LUSK, G , Am. Jour. Physiol., **16**, 129 (1906).
- (6) KONDO, K , Biochem. Zeit., **45**, 63 (1912); Noorden, K. v., Biochem. Zeit., **45**, 94 (1912).
 LEVENE, P. A , AND MEYER, G M , Jour. Biol. Chem., **11**, 361 (1912).
 SLOSSE, A , Arch. intern. de Physiol , **11**, 154 (1912).
- (7) LEVENE, P A , AND MEYER, G. M , Jour. Biol. Chem , **14**, 149 (1913).
- (8) DAKIN, H D., AND DUDLEY, H W , Jour. Biol. Chem., **14**, 155 (1913).
 LEVENE, P A , AND MEYER, G M , Jour. Biol. Chem , **14**, 551 (1913).
 NEUBERG, C , Biochem. Zeit , **49**, 502 (1913).
- (9) BUCHNER, E , AND MEISENHEIMER, J , Ber. deut. Chem. Ges., **39**, 3201 (1906); **43**, 1773 (1910).
 MAYER, P , Biochem. Zeit , **2**, 435 (1907).
 WOHL, A , Biochem. Zeit , **5**, 45 (1907).
 LEBEDEV, A , Biochem. J , **11**, 189 (1917).
- (10) BERGMANN, K , Ann. Chem , **436**, 173 (1924).
- (11) LEVENE, P. A , AND WALTI, A , Jour. Biol. Chem **73**, 263 (1927).
- (12) LEVENE, P A., AND WALTI, A., Jour. Biol. Chem., **75**, 325 (1927).
- (13) EULER, H v , AND LUNDEQVIST, G , Zeit. physiol., Chem , **72**, 97 (1911).
 WILLSTATTER, R , AND SOBOTKA, H , Zeit. physiol. Chem , **123**, 170 (1922).
- (14) FERNBACH, A , AND SCHILLER, N , Compt. rend , **178**, 2196 (1924).
- (15) WILLSTATTER, R , and SOBOTKA, H , Zeit. physiol. Chem , **123**, 170 (1922).
- (16) HARDEN, A , AND YOUNG, W J , Proc. Roy. Soc. B, **81**, 336 (1909).
 See also Herzog, R O , and Saladin, O , Zeit. physiol. Chem., **73**, 263 (1911).
- (17) LEVENE, P A , AND MEYER, G M , Jour. Biol. Chem , **14**, 149 (1913); **15**, 65 (1913).
- (18) LAQUER, F , Zeit. physiol. Chem , **116**, 169 (1921); **122**, 26 (1922).
 LAQUER, F , AND MEYER, P , Zeit. physiol. Chem , **124**, 211 (1923).
- (19) WARBURG, O , POSENER, K , AND NIEGELIN E , Biochem. Zeit., **152**, 309 (1924).
- (20) EULER, H v , MYRBACK, K , AND KARLSSON, S , Zeit. physiol. Chem., **143**, 243 (1925).
 EULER, H v., AND NILSSON, R , Zeit. physiol. Chem , **148**, 211 (1925).
- (21) EMBDEN, G , GRIESBACH, W., AND SCHMITZ, E., Zeit. physiol. Chem., **93**, 1 (1914).
 EMBDEN, G., GRIESBACH, W., AND LAQUER, F., Zeit. physiol. Chem., **93**, 124 (1914).
 EMBDEN, G., AND HAYMANN, C., Zeit. physiol. Chem , **137**, 154 (1924).
 EMBDEN, G , AND ZIMMERMANN, M , Zeit. physiol. Chem , **141**, 225 (1924).

- (22) NEUBERG, C , AND GOTTSCHALK, A., *Biochem Zeit* , **146**, 164 (1924).
- (23) LEBEDEV, A , *Biochem Zeit* , **10**, 454 (1908); **28**, 213 (1910).
EULER, H. v., AND BERGGREN, T , *Zeit f Garungsphysiologie*, **1**, 203, (1912).
EULER, H v., AND JOHANSSON, D., *Zeit physiol Chem* , **76**, 347 (1912).
EULER, H v , *Zeit. physiol Chem.*, **89**, 337, (1914); **90**, 355 (1914)
- (24) HEWITT, J A , AND PRYDE, J , *Biochem Jour* , **14**, 395 (1920)
WINTER, L B., AND SMITH, W , *Jour. physiol* , **57**, 100 (1923); **58**, 327 (1924).
LUNDGAARD, C , AND HOLBØLL, S , *Jour Biol. Chem* , **62**, 453 (1924), etc.
- (25) PRYDE, J , *J Chem Soc.*, **123**, 1808 (1923)
LEVENE, P A , AND MEYER, G M , *Jour Biol Chem* , **60**, 167 (1924).
CHARLTON, W., HAWORTH, W. N , AND PEAT, S , *J Chem. Soc.*, **1926**, 89.
LEVENE, P A , AND SIMMS, H S , *J Biol Chem* , **68**, 737 (1926).
- (26) LEVENE, P A , AND SIMMS, H S , *Jour Biol Chem* , **65**, 31 (1925).
- (27) LEVENE, P A , AND MEYER, G M , *J Biol Chem* , **74**, 701 (1927).
- (28) WILLSTATTER, R , AND SOBOTKA, H , *Zeit physiol Chem* , **123**, 164 (1922).

THE MINERALOGY OF HAFNIUM*

O. IVAN LEE

Process Laboratories, Inc , Irvington, New Jersey

Although it has been but a scant five years since the advent of hafnium (1), the missing element of atomic number 72, a considerable literature has accumulated, one of the most complete accounts of the element being embodied in the recent monograph by the discoverer, Georg v. Hevesy (44). The papers published, however, have dealt almost exclusively with the chemistry and physics of the metal, and little has been said about the minerals in which it is found. As there are indications that hafnium is emerging from the laboratory to the ranks of the commercial rarer metals, some account of this rather neglected phase of the subject is of general interest and may be of assistance in making the metal more accessible. Inasmuch as elements closely related chemically are often coupled in minerals, as, for example, copper and silver, zinc and cadmium, gallium and indium, germanium and tin, columbium and tantalum, molybdenum and tungsten, and the rare earths, it will be instructive briefly to review the salient facts which led to the association of hafnium with zirconium and its discovery therewith.

THE RARE EARTHS AND THEIR RELATION TO HAFNIA

Ever since Mendelejeff and Lothar Meyer conceived their periodic arrangement of the elements, certain discrepancies were known but ignored for the sake of the more fundamental harmony. The magnetic triad of iron, cobalt and nickel, and the noble ones of the platinum metals were disturbing, but even more so were that significant group of numerous metals follow-

* Based on a paper presented to the New York Section of the American Chemical Society, December 10, 1928.

ing barium, known collectively as the rare earths. So far as then known, other triads could follow and new rare earths might

THREE OUTER
(VALENCE)
ELECTRONS

FOUR OUTER
(VALENCE)
ELECTRONS

C⁶
(12.00)

Si¹⁴
(28.08)

Ti²²
(48.1)

Ge³²
(72.60)

FIVE ELECTRONIC
ORBITS

Zr⁴⁰
(91.2)

Sn⁵⁰
(118.7)

SIX ELECTRONIC
ORBITS

Cs⁵⁵
(132.9)

Ba⁵⁶
(137.4)

* 57-71
Rare Earths
(138.9 - 175.0)

Hf⁷²
(178.6)

Ta⁷³
(181.5)

W⁷⁴
(184.0)

Th⁹⁰
(232.1)

Pb⁸²
(207.2)

* 57-71
Rare Earths
(138.9 - 175.0)

La⁵⁷
(138.9)

Ce⁵⁸
(140.2)

Pr⁵⁹
(140.9)

Nd⁶⁰
(144.3)

Il⁶¹
(144.6)

Sm⁶²
(150.4)

Eu⁶³
(152.0)

Gd⁶⁴
(157.3)

Tb⁶⁵
(159.2)

Ds⁶⁶
(162.5)

Ho⁶⁷
(163.5)

Er⁶⁸
(167.7)

Tm⁶⁹
(169.4)

Yb⁷⁰
(173.5)

Lu⁷¹
(175.0)

FIG 1 HAFNIUM AND RELATED ELEMENTS ARRANGED ACCORDING TO MOSELEY'S
LAW OF ATOMIC NUMBERS¹

(Atomic weights² given in parentheses)

be sought; in a word, no valid limitation could be placed on the number of elements possible, and it might even be greater

¹ After H. D. Hubbard, U. S. Bureau of Standards, Periodic Chart of the Atoms, Pub. W. M. Welch Mfg. Co., Chicago, Ill., 1926

² Berichte 60, 1; 1 (1927), II excepted

than could be accommodated by the periodic system. Mendelejeff estimated the rare earths as numbering twenty. But Moseley's concept of atomic numbers has demonstrated that there is room for but fifteen, all now known since Hopkins' discovery (2) of element number 61, illinium.

Of the fourteen rare earths known at the time of Moseley's postulation, those with atomic number 70 and 71 were the last discovered. In 1878, Marignac isolated a new earth from the old erbia, slightly more basic, and which was named ytterbium. In 1905, Welsbach found that ytterbium was a mixture of two elements, and in 1907 named them aldebaranium and cassiopeium, and they have since been identified as elements number 70 and 71. Urbain, however, had already confirmed their entity and named them neoytterbium and lutecium, and the names finally adopted in 1914 were ytterbium and lutecium. It was suspected by both Welsbach and Urbain that Marignac's ytterbium contained yet a third element, and both investigators carefully searched for it. Welsbach failed to find any, but Urbain was more successful, and in 1911 announced the discovery of this new rare earth which he called celtium. In 1922, Dauvillier stated that Urbain's preparation showed faint traces of elements number 69 and 72 in addition, indicating in 1923 that element 72 was present in the proportion of about 1:10,000. Meanwhile, Hevesy had taken up the question, and in 1922, by virtue of the quantum theory of atomic structure (3) asserted that element 72 should indeed be a constituent of an "old element," but of "zirconium" and not of "ytterbium" (fig. 1).

PROBABLE ABUNDANCE OF THE MISSING ELEMENT, AND RESEMBLANCE TO ZIRCONIUM

It could not be foretold by the pioneer investigators whether element No. 72 was abundant enough in a mineral to be recognized in the specimens to be examined, but there was good reason to believe that the element sought was not rare. For instance, it was known that elements with odd atomic numbers, such as gallium, 31, indium, 49, and thallium, 81, are much more scarce than those with even, such as germanium, 32, tin, 50,

and lead, 82. A study of the relative abundance of silica, titanium, zirconium and thorium in the lithosphere was also reassuring on this point, their oxides being found in the earth's crust in the percentages respectively of 59.09, 1.05, 0.04 and 0.002.

Bohr's theory of atomic structure (3) demanded that element 72 must differ entirely in its outer structure from that possessed by the rare earths which precede it, so that the difference between lutecium and element 72 is greater than that of any two of the rare earths. In addition, its atomic structure shows a very marked analogy to that of element number 40, zirconium. The simultaneous addition of another electron with one more in the outer ring causes element 72 to be quadrivalent. All these considerations gave weight to the belief that element 72 must be analogous to zirconium and therefore likely to be found with it, thus leading to the examination of zirconium minerals and preparations in which the new element was soon identified and isolated³. This association is so characteristic that as yet hafnium has not been found other than with the older element discovered by Klaproth in 1789.

The chemical similarity of the two metals is so great that their separation is attended with difficulty, for *hafnium and zirconium exhibit the same chemical reactions*. The fractionation of the hexafluorides makes possible a separation which, while much more difficult than that of columbium and tantalum, is nevertheless easier than most rare earth separations (4).

SUGGESTED USE OF CLERICI SOLUTION FOR THE CONTROL OF HIGHER FRACTIONATION

In connection with the resolution of the fluorides of hafnium and zirconium, it occurred to the writer that the little known Clerici solution (5) (6) (7) (8), a molecular mixture of thallium malonate and thallium formate, might be used for testing the

³ A small sample of the first metallic hafnium ever made, presented by G. Hevesy, the discoverer, is on permanent exhibition at the American Museum of Natural History, New York, in the Element Collection being compiled by G. F. Kunz, Curator of Research

ammonium hafnium-zirconium hexafluoride crystals with densities too high to float on even fortified methylene iodide, if not soluble or incompatible.

Through the generosity of C. Palache of the Department of Mineralogy, Harvard University, some Clerici solution was placed at the author's disposal, and it was ascertained that ammonium hexafluozirconate floats readily in the diluted solution without sensible solution or decomposition. The only ammonium fluohafniate available appeared in tetragonal instead of monoclinic form, and was possibly heptafluoride. However, it sank in methylene iodide but less readily in the dilute Clerici solution. The finer crystals slowly dissolved, but the coarser ones remained a long time, ample for observation, so that it is believed that this mineralogical reagent should prove useful in advanced hafnium fractionation.

THE OCCURRENCE OF HAFNIUM IN NATURE

It will be recalled that in the case of chlorine, no matter from what source it may have been derived, the atomic weight is always the same in spite of the fact that it is a mixture of two isotopes of atomic weights 35 and 37. All the chemical vicissitudes of eons have not disturbed the ratio which was presumably present in the primeval homogeneous mixture. On the other hand, since zirconium and hafnium are chemically separable, some change should have taken place in their original ratio, reflecting in various minerals different geochemical processes in the earth's history. It has been demonstrated that this difference in the ratio of HfO_2 to ZrO_2 is a reality, and ranges from 0.007 for the bean-like favas of Brazil to 0.5 for the rare scandium mineral "befanamite" of Madagascar. In general, minerals of granitic origin exhibit a decidedly higher ratio, but the ratios are never so marked as those of less dissimilar elements occurring in pairs such as columbium and tantalum or molybdenum and tungsten. The association of Hf and Zr and other paired elements finds an adequate explanation in the studies of Goldschmidt (9) which support his belief that the geochemical distribution of the elements is conditioned by the electron shell.

TABLE I
List of zirconium minerals known to contain hafnium
(Specific paired analyses are italicized)

NAME	FORMULA	Per cent ZrO ₂	Per cent HfO ₂
Baddeleyite (distinct crystals)	ZrO ₂	96.5-97.7-98.9	1.2
Syn. Brazilite (fibrous, botryoidal, columnar) ..	ZrO ₂	71-93	
Syn. var (rock) Jacupirangite (crystallized)	ZrO ₂	59-74-92.4	0.5-0.5-0.7
Zirconia ore	ZrO ₂	71-93	
	ZrO ₂ SiO ₂	67 80-85	
1 Favas (alluvial pebbles, Brazilite)	?	75	
2 "Zirkite" (mixture)	(Na ₂ , Ca)O ZrO ₂ 3SiO ₂ 2H ₂ O	31.53	0.3
Catapleite	Na ₂ O·ZrO ₂ 6SiO ₂ ·3H ₂ O	20.28	0.2
	Na ₁₃ (CaFe) ₄ Cl SiZr ₂₀ O ₃₂	12.20-14.32	0.1-0.2-0.17
Elpidite	Na ₁₃ (CaFe) ₄ Cl SiZr ₂₀ O ₃₂	12.21	0.2
Eudalyte	5RTiO ₃ 3RZrO ₃ R(NbTa) ₂ O ₆ (R = Ce, Fe, Ca)	29.11	0.6
Var Eucolite	CaNi ₂ O ₆ NaF	2.90	tr.
Polymignite	6CaSiO ₃ 2Na ₂ ZrO ₂ F ₂ (Ti·SiO ₃) ₂ (TiO ₃)	19.80	0.3
Pyrochlore	(ScYt) ₂ Si ₂ O ₇	2	0.5
Rosenbuschite	Sc ₂ Si ₂ O ₇	1.3	1.0
Thortveitite	13(CaNa ₂)O 9SiO ₂ 3ZrO ₂ ·Nb ₂ O ₅	15.61	0.5
Var "Befanamite"	ZrO ₂ ·SiO ₂	64.23	0.98*
Wöhlerite	ZrO ₂ ·SiO ₂	64.83	1.2
Zircon	(Zr, Hf, Th, Be)O ₂ SiO ₂ nH ₂ O	41.98	4.6
Var Hyacinth	R ₂ Y ₂ (Zr, Hf) (SiO ₄) ₁₂	52.4	5.5
Alt.	3·SiO ₂ Zr, HfO ₂ H ₂ O	53.2-65.18	3.4-2.6
Malakon	ZrO ₂ SiO ₂ with Yt, Th, Nb and Ta	49.8	3.5
Naegeite	(CaFe)O 2(ZrTiTh)O ₂	51.89	1
Zirkelite			

* Mean value from six analyses

The hafnium content of virtually every known type of zirconium mineral has been determined, and it has developed that *in the case of zircons, the amount of hafnium is roughly proportional to the radioactivity* (10). The reason for this relation is interesting. The radioactivity is due to uranium and thorium, and hafnium is isomorphous with these elements (11). E. and G. Urbain (12) have discussed the presence of both hafnium and the yttrium elements in some zirconium minerals, and find that the more hafnium is present the larger the proportion of rare earths. Strangely enough, however, hafnium has not been found in any thorium mineral such as thorite or thorianite, nor in titanium minerals like rutile and ilmenite.

THE HAFNIUM CONTENT OF THE EARTH'S CRUST

If the average hafnium content of zirconium is assumed to be 3 per cent, the hafnium content of the earth's crust works out to be 1:100,000. Fortunately, nature has made the search easier by segregating the element in certain minerals, among which may be mentioned the following species, varieties and alterations in which hafnium has actually been determined (table 1).

Although numerous zirconium minerals have been investigated for their hafnium content, the field is by no means exhausted inasmuch as the species and varieties enumerated in table 2 are not specifically mentioned as having been tested. Most of these are very rare, and the status of some, such as the varieties of zircon, is not well defined; nevertheless it is of interest and value to test doubtful and even discredited species and varieties as an additional means of checking their validity.

THE OCCURRENCE OF HAFNIUM MINERALS IN THE UNITED STATES

Only two zirconium minerals have ever been systematically mined, the ortho silicate, zircon, and in more recent years, the native oxide, brazilite. Only the former occurs in this country and in commercial quantity in but three localities. Zircon was mined intermittently in western North Carolina from 1869 to 1911 (13), as much as 26 tons having been taken out in 1883 by

TABLE 2
List of zirconium minerals not reported investigated for hafnium

NAME	FORMULA	PERCENT OF ZrO ₂
Astrophyllite	(KNaH) ₉ FeMnCaMg ₁₉ FeAl ₁₁ (TiZr)(OF ₂) ₃ [SiO ₄]	1 21-4 97
Beckelite	Ca ₃ (YtCeLaNdPr ₁₄ SiZr) ₃ O ₁₃	2 5
Chalcoclamprite	R ^{II} Nb ₂ O ₆ F ₂ R ^{II} SiO ₃ (R = Ce, Na, Zr and Ca) (allied to pyrochlore)	5 71
Enderohite	R ^{II} Nb ₂ O ₆ (OH) ₂ R ^{II} SiO ₃ R ^{II} = Na, Ca; R ^{III} = Ce	3 78
Guarinite	3CaSiO ₃ (CaFOH) ₂ NaZrO ₃	19 70
Hortdahlite	6Ca ₂ Si ₂ O ₆ 2[(Na ₂ CaH ₂) ₂ Zr ₂ F ₄ O ₄]/2 3[CaMgFeMn](ZrTiSi ₂) ₂ O ₆	21 48
Johnstrupite	Na ₃ Ca ₃ Ce ₂ FOH ₁₇ SiO ₃ ¹⁰	2 84
Lävenite	Na[ZrOF](MnCaFe)(SiO ₃) ₂	28 79-28 90
Leucosphenite	BaSi ₄ O ₉ 2Na ₂ (TiZr) ₂ Si ₂ O ₉	3 5
Loranskite (var. wilkite)	Tantalate of Yt, Ce, Zr, Fe, etc	20 00
Lorenzenite (ramsayite)	Na ₂ (TiZr) ₂ O ₉ Na ₂ Si ₄ O ₉	11 92
Mosandrite	Very near johnstrupite	7 43
Nohlite (var samarskite)	Hydrated niobate of U, Yt and Fe	2 96
Oliveirite (alt euxenite)	3ZrO ₂ 2TiO ₂ 2H ₂ O	63 36
Orvillite (in zircon in "caldasite")	8ZrO ₂ 6SiO ₂ 5H ₂ O	68 04
Riebeckite (arfvedsonite)	4Na ₂ SiO ₃ 5FeSiO ₃ 5Fe ₃ Si ₃ O ₉	0 75-4 7
Soda-catapleite (var catapleite)	Na ₂ O ZrO ₂ 3SiO ₂ 2H ₂ O	30 80
Uhlignite	3 3Ca(TiZr) ₂ O ₆ ·Al ₂ TiO ₃	21 95
Zircon	ZrO ₂ ·SiO ₂	38-61 53-69
Alt. Auerbachite		42 0
Var "Hagatallite"		40 9
Var "Oxymalite"		

W. E. Hidden (14) More recently, the Pablo Beach deposit of rare minerals on the East Coast of Florida has yielded 2 per cent of zircon sand as an adjunct to the electromagnetic separation of garnet, monazite, rutile and ilmenite. The zircon is stated to swell greatly when ground, disengaging large quantities of helium (15). The occurrence of a rich zirconiferous sandstone near Ashland, Va., has also been described (16). Unfortunately unaltered zircon does not appear to be a promising source of hafnium as the content is usually less than two per cent of HfO_2 , although a specimen from Henderson County, N. C., is stated to contain more (17). Of the altered zircons, cyrtolite is by far the most common and widely distributed, having been reported from at least seven states—Mt. Antero, Colorado, Branchville, Conn., Rockport, Mass., Mitchell and Henderson Counties, N. C. (three localities) (18), Westchester County, N. Y., Pennsylvania (five localities) (19) and Llano County, Texas, and the Massachusetts and New York minerals have already been shown to be among the richest hafnium minerals known. In addition, a hydrated zircon from Amherst County, Virginia, first described by Mallett (20) is under examination.

The cyrtolite from Texas does not appear to have been analyzed for hafnium as yet, but since it is known to have given "good radiographs after twenty-four hours exposure" (21) it is a fair presumption that the mineral should prove a rich source of hafnia. Hidden and Mackintosh (22) state that it was abundant at the famous Banger Hill locality during the mining of the yttria minerals, many hundred pounds having been found in a great variety of form and color, but this mine is no longer operated.

The cyrtolite from Rockport, Mass., is high in hafnia (9 to 17 per cent (?)), but very rare, and that from Connecticut and Colorado is also very sparing in occurrence, but the radioactive cyrtolite from Westchester County, New York, has twice been authenticated as high in hafnia and is fully as abundant as the Texas deposit.

THE OCCURRENCE OF HAFNIUM NEAR NEW YORK CITY

Up in the hills of Westchester County, near the historic village of Bedford, is a large feldspar mine which has long been a Mecca

for mineralogists from the metropolitan district, since the great pegmatite vein to which the quarry owes its origin, is a prolific and unfailing source of a variety of those minerals which delight the collector. Among these may be mentioned asteriated rose quartz, beryl, columbite, garnet and black tourmaline.

Early in the history of the mine when the workings were not so deep, an unusual, heavy, dark brownish-black mineral of adamantine luster altering to brownish-red or chocolate color of duller aspect, was noticed. It occurred in pockets or isolated



FIG. 2. CARTOLITE (ALTERED ZIRCON) RICH IN HAFNIA, FROM NEAR BEDFORD, WESTCHESTER COUNTY, NEW YORK

Characteristic curved tetragonal pyramidal crystals are shown penetrating feldspar. This mineral was used by C. James of the University of New Hampshire as the source of the first hafnium and pure zirconium salts prepared in this country (1924). Hafnia 5.5 per cent (9.5 per cent of $\text{ZrO}_2 + \text{HfO}_2$) Zirconia 52.4 per cent. Natural size.

vugs in contact with both smoky quartz and orthoclase, intimately associated with yellowish secondary uranium minerals, one identified as autunite. The presence of these, coupled with the black color, unusual weight and pitchy luster, and tests for uranium, led to the mineral being mistaken for pitchblende or uraninite, then attracting considerable interest because of the recent discovery of radium.

L. V. Case of Tarrytown, N. Y., the original finder, submitted specimens to the department of mineralogy of Columbia University, New York, and Luquer (23) published a paper

showing that the strange mineral was not pitchblende, but the altered zircon known as cyrtolite. It is characterized and named because of the curved form of its crystals, which are tetragonal pyramids not infrequently resembling rhombic dodecahedrons (fig. 2). The unaltered black portions of higher density were found to contain minute lacunae of uraninite, thus accounting for the uranium tests and alteration products, one of which was identified as uraconite, a mineral new to this locality. Striking radiographs are easily made with it (fig. 3), a fact of



FIG. 3 RADIOGRAPH OF HIGHLY HAFNIFEROUS CYRTOLITE, NEAR BEDFORD, WESTCHESTER COUNTY, NEW YORK

Hafnia. 5.5 per cent (9.5 per cent of $ZrO_2 + HfO_2$) Zirconia. 52.4 per cent. Exposure. 22 days on bare plate. Courtesy of T. I. Miller, Newark Mineralogical Society, Newark, N. J.

especial practical significance to the prospector and mining engineer. A rough analysis of the cyrtolite showed 27.24 per cent of silica, 53.56 per cent of zirconium oxide, and 1.14 to 4.35 per cent of uranium oxide, together with traces of copper, calcium, manganese, "*and probably of yttrium.*" An exact analysis was not considered advisable because of the altered character of the mineral.

About four years ago, the author recalling the statement italicized above, suggested to C. James of the University of New

Hampshire, that the suspected yttria might warrant detailed study, a fifteen pound sample of the mineral subsequently being forwarded. An investigation by him revealed that the presence of yttria, erbia, thulia, etc. was transcended by a notable percentage of the new element hafnium, approximately 10 per cent of hafnium oxide being present in the mixed $\text{HfO}_2 + \text{ZrO}_2$ as obtained

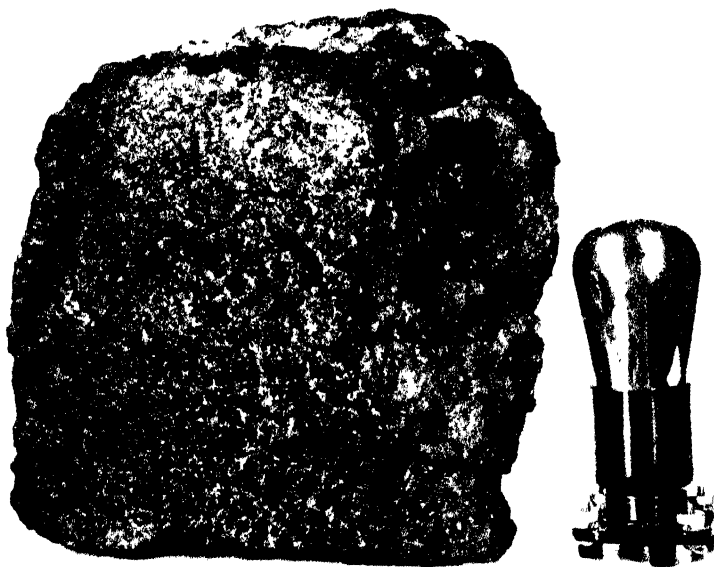


FIG. 4. LARGE SPECIMEN OF HIGHLY HAFNIFEROUS CYRTOHITE, NEAR BEDFORD, WESTCHESTER COUNTY, NEW YORK

Weight 7.5 kgm (16.5 pounds) Hafnia 0.41 kgm (0.91 pounds) (A standard radio tube is shown for comparison.)

from the mineral (24). In a later letter from Hevesy (25), James' original analysis was confirmed in that a preliminary analysis "very near to the final value" showed 9.5 per cent of hafnium oxide in the mixed zirconium-hafnium oxides which total 52.44 per cent. Unlike the Norwegian alvite, this unusual percentage of the New York mineral appears to be rather consistently maintained, although the yttrium earths are very variable, and a large

quantity was afterwards successfully worked up by James for hafnium and the concomitant zirconium. This is believed to have constituted the first production of these elements in this country.⁴ With the possible exception of the Rockport cyrtolite⁵ the mineral occurring at Bedford is apparently the richest ore of hafnium yet reported (44). Although the deposit is of limited extent, specimens weighing ten or twelve pounds are not uncommon, and one is illustrated which perhaps contains more hafnium for its size than any other extant (fig. 4).

FOREIGN DEPOSITS OF HAFNIUM

Abroad, the principal commercial occurrences of zircon are in the monazite sands of Travancore, India, and of Brazil, and in the pegmatites of Southern Norway, but the mineral has largely lost its former importance as a source of zirconia because of the discovery of the native oxide, baddeleyite. The pure crystallized mineral has been known for thirty-five years, and has been reported from Ceylon, Italy, Sweden and Montana, but it remained for the Brazilian states of Minas Geraes and São Paulo to disclose to the world in 1915, zirconia ore in unprecedented quantity (26). The ore (fig. 5) outcrops on a mountainous plateau, and the methods of mining and transportation employed are very primitive owing to the distance from railroads and the rugged character of the region. Isolated boulders have been found weighing as much as $27\frac{1}{4}$ milliers (thirty tons), so that such a "specimen" would presumably contain from 136 to 190 kgm. (300 to 420 pounds) of hafnium oxide.

A third hafnium mineral of considerable interest is the com-

⁴ Specimens from G. Hevesy of the first hafnium and pure zirconium salts ever isolated, and of the first made in America, from C. James, may be seen in the Chandler Chemical Museum of Columbia University, New York, with related exhibits of historic interest, including Bedford cyrtolite from which the American salts were produced, and a cast lead dish used in their fractionation.

⁵ A reprint of "Recherches sur les propriétés du hafnium" (København, 1925) received by the author from G. Hevesy contained the marginal annotation against the hafnia content of 17 per cent reported for analysis 23 (Cyrtolite, Rockport, Mass., U. S. A.), "Uncertain! Must be repeated." This is the same mineral previously reported (G. Hevesy and V. T. Jantzen, *J. Chem. Soc.* **123**, 3218 (1923), analysis 20) as giving 9 per cent of HfO_2 .

plex eudialyte (and var. eucolite) reported in the United States only from Arkansas (27) and whose occurrence in Hibina-Toundra and Lujavr-urt in the Kola Peninsula of Russian Lapland was recently described by Fersman (28).

In spite of the trying nature of the district--1200 meters above sea level, covered with forests and swamps, infested with insects, devoid of roads, paths or habitations, with snow falling as early as August 9th--the Mineralogical Museum of the Academy of Sciences of the USSR, and the Research Institute of the North have sent six expeditions to this inhospitable country to collect the rare and unique minerals found there in



FIG. 5 LOW GRADE HAFNIA ORE ABUNDANT IN BRAZIL. FIBROUS AND BOTRYOIDAL BRAZILITE (BADDELEYITE), CALDAS REGION, STATES OF MINAS GERAES AND SÃO PAULO

Hafnia 0.5-1.2 per cent Zirconia 71-93 per cent Natural size

profusion. Over five tons of specimens have been collected and are now in part displayed in Leningrad. Among them are about a dozen of the rarer zirconium minerals. Only the eudialyte has been analyzed for hafnium as yet so that this field is a virgin one.

Gordon (29), however, in his description of the noted cryolite mine at Ivigtut, Greenland, notes that eudialyte is of common occurrence in the alkaline syenites about the Tunugdliarfik and Kangerdluarsuk fiords of the Julianehaab district; and 55 tons were collected for technical purposes according to K. J. V. Steenstrup (30). It occurs disseminated in the rock as well as

in veins, and Gordon remarks, perhaps prophetically, that "this district should become an important producer of this mineral if a use for it is ever discovered." A typical specimen of this rare species from this remote locality is illustrated in Figure 6.

It is unfortunate that all of the three minerals aforementioned which are known to be abundant at favored but inaccessible localities, should be so lean in hafnia, even zircon seldom exceeding 1.2 per cent. In this connection, it is well to remember that James (31) has pointed out that *so long as one has a use for*



FIG. 6. A RARE AND COMPLEX HAFNIUM MINERAL COMMON IN SOUTHERN GREENLAND. TYPICAL CRYSTALLIZED EUDIALYTE FROM KANGERDLUARSUK.

Hafnia 0.17 per cent. Zirconia: 14.3 per cent. Natural size.

pure zirconium material, zirconia of very low hafnia concentration can be utilized. From a strictly scientific viewpoint, the properties of pure zirconium, contaminated as it was for one hundred and thirty-four years, should claim equal attention with those of the offending hafnium, known for only five. Nevertheless, since commercial and historical zirconium preparations examined by Hevesy and others contained only small percentages of hafnium, the general properties already chronicled for the older element do not exhibit marked divergences from those found for

TABLE 1
The occurrence of hafnium minerals by countries

LOCALITY	MINERAL	ZrO ₂	HfO ₂
<i>Africa</i>			
Diego Suarez	Zircon		0.8
Madagascar	Zircon		0.9
Madagascar	Grey zircon		0.8
Madagascar, Befanamo	Befanamite (thortveitite)	1.3	1.0
Madagascar	Malakon (alt. zircon)	53.2	4
<i>Asia</i>			
Ceylon	Becarite (var. zircon)		2.1
Ceylon	Reddish-brown zircon		2.7
Ceylon	Zirkelite	51.89	1
India (Travancore?)	Zircon from monazite	64.0	1.2
Japan, Mino	Naegite	48.30	7
Japan, Naegi	Naegite	49.8	3.5
Siam, Province of Chantaboon	Zircon (transparent blue)		4.2
<i>Australia</i> No minerals have yet been definitely reported from the continent as containing hafnium			
Tasmania	Zircon		*
Tasmania	Greyish brown zircon		1.1
<i>Europe</i>			
Austria, Carinthia	Zircon (transparent white)		4.2
France, Espailly (Le Puy)	Zircon (var. hyacinth; transparent red)	64.83	1.2
France, Espailly (Le Puy)	Zircon (var. hyacinth; transparent red)		0.7
France, Espailly (Le Puy)	Zircon (transparent green)		1.1
Italy, Lonedo	Zircon (yellow)		0.7
Italy, Lonedo	Zircon (red)		0.7

Italy, Vicenza...	Zircon (transparent green)		0.8
Italy, Vesuvius	Zircon		0.7
Norway, Barkevik	Eudialyte (var. eucolite)	14 47	0.7
Norway, Barkevik	Eudialyte (var. eucolite)	12 21	0.2
Norway, Barkevik	Wöhlerite	15 61	0.5
Norway, Brevik	Zircon		1.0
Norway, Brevik	Reddish-brown zircon	63.2	1.0
Norway, Frederiksvärn	Polymignite	29.11	0.6
Norway, Frederiksvärn	Brown zircon	65 2	1.0
Norway, Gjersted	Alvite (alt. zircon)		9
Norway, Iveland	Thortveitite	2	0.5
Norway, Hitterö	Malaccon (alt. zircon)	65 18	2.6
Norway, Kragerö	Alvite (alt. zircon)	41 98	4.6
Norway, Kragerö	Alvite (alt. zircon)		3.
Norway, Kragerö	Alvite (alt. zircon)		8.
Norway, Kragerö	Alvite (alt. zircon)		15.
Norway,	Catapleite	31 52	0.3
Norway, Langesund	Brown zircon		1.7
Norway, Langesund	Rosenbuschite	19.80	0.3
Norway, Larvik	Greyish brown zircon		6.
Norway, Risør	Alvite (alt. Zircon)		10.
Norway	Grey syenite (rock)		3.8
Norway, Unneland	Thortveitite	0.8	1.1
Sweden, Alno	Pyrochlore	2 90	tr.
Russia, Miask	Greyish brown zircon	64 22	1.1
Russia, Rojokw Kliutsch, Ural	Brown zircon		0.5
Russia, Kola	Eudialyte (red)		0.1
<i>North America:</i>			
Canada, Eganville, Ont.	Brown zircon		1.2
Canada, Renfrew County, Ont	Zircon		0.6

TABLE 3—Continued

LOCALITY	MINERAL	ZrO ₂	HfO ₂
<i>North America—Continued:</i>			
Greenland	Zircon		0.8
Greenland	Catapleite	31.53	0.3
Greenland, Narsarsuk	Elpidite	20.28	0.2
Greenland, Narsarsuk	Reddish-brown zircon		0.8
Greenland, Narsarsuk	Eudialyte (red)	12.20	0.2
Greenland, Narsarsuk	Eudialyte (brown)	12-16	0.6
Greenland, Kangerdluarsuk	Eudialyte	14.32	0.17
United States, Connecticut	Zircon		1.0
United States, Rockport, Massachusetts	Cyrtolite (alt. zircon)	40	9
United States, Rockport, Massachusetts	Cyrtolite (alt. zircon)	44	17 (?)
United States, Bedford, New York	Cyrtolite (alt. zircon)	52.4	5.5
United States, North Carolina	Grey zircon		4 (?)
United States, North Carolina	Brown zircon		1.3
County	Zircon		4.0 (?)
<i>Oceania:</i> No minerals have yet been reported definitely as containing hafnium.			
<i>South America:</i>			
Brazil	Baddeleyite	97.7	1.2
Brazil	Fava	92.42	0.7
Brazil	Fava (shell)	59	0.5
Brazil	Fava (nucleus)	74	0.5
Brazil	Zircon separated from monazite sand	64	0.4
Brazil, Caldas	Zircon		1.8
Brazil, Minas Geraes	Zircon		1.0

* "Very considerable."

hafnium-free zirconium preparations. On the other hand, the properties of hafnium are very remarkably different from those of zirconium, so that interest has naturally centered in the newer element and the minerals richest in it.

Of the altered zircons, cyrtolite is known in India, Madagascar, Italy, Sweden and Hybla, Ontario, and Norwegian alvite has been useful to the early investigators of hafnium, but does not seem to be a common mineral.

For the guidance of those interested, table 3 presents the distribution of zirconium minerals whose hafnium content has been definitely determined (32) but the list can ultimately be made far more comprehensive as more analyses accumulate.

It is to be feared that the prospector and the mining engineer will be compelled for some time to come to rely on well equipped chemists of skill and experience for authentic information concerning the actual hafnium content of suspected minerals and ores. Meanwhile, it may be useful to point out that since the various types of altered zircon which have proven richest in the new element, are pseudomorphic after zircon, it is not difficult to become accustomed to recognizing their square pyramidal form, particularly if specimens of the various alterations of zircon are studied. If such a mineral is identified and found to be in addition markedly radioactive, it is reasonably certain that the mineral under examination will prove to contain an appreciable percentage of hafnia.

CONCLUSION

In conclusion, it must not be forgotten that the discovery of hafnium has necessitated the revision and correction of all the data on zirconium so painfully acquired over the course of a century and a third. The International Critical Tables warns that "All zirconium salts (are) probably contaminated with 1-5 per cent of hafnium" (33). So far as hafnium is concerned, it is somewhat early to ask for what practical purpose it may ultimately prove singularly appropriate, for the process of accumulating and studying facts about it is still actively engaging the attention of many chemists and physicists. An excellent

summary of the commercial uses of zirconium set forth by Meyer (34) will be found very suggestive, however, to those of a pragmatic turn of mind. It may be of interest also, to note that the close similarity of metallic hafnium to thorium coupled with its high melting point and electronic emissivity (35), have already led to patents being taken out for its use in lamp filaments, particularly those used in radio tubes (36).

The author extends his sincere thanks to G. Hevesy of the Physikalischchemisches Institut of the University of Freiburg, Germany, and to C. James of the Department of Chemistry of the University of New Hampshire, Durham, N. H., for their cordiality and helpful coöperation in supplying numerous items of interest embodied in this article.

REFERENCES

- (1) COSTER, D., AND HEVESY, G. *Nature* **111**, 79, 182, 252, 462 (1923).
- (2) HARRIS, J. A., AND HOPKINS, B. S.: *J. Am. Chem. Soc.* **48**, 1585-94 (1926).
HARRIS, J. A., YNTEMA, L. F., AND HOPKINS, B. S. *Ibid.* 1594-8.
- (3) BOHR, N.: *Theory of spectra and atomic constitution*, Cambridge University Press, 2nd edition (1924).
- (4) JAMES, C.: *Private communication*, Oct. 20, 1926.
- (5) CLERICI, E. *Atti accad. Lincei* **16**, 187-195 (1907)
- (6) ESKOLA, P.: *Videnskapsselskapets-Skrifter. I Matemat. natur Klasse (Kristiania)* **8**, 6 (1921).
- (7) HORNER, F.: *Beiträge zur Kenntnis des Staurolits*. Inaug. Diss. Heidelberg **1915**, quoted in Rosenbusch-Wülfing, *Mik. Physiographie* I, p. 692 (1924).
- (8) VASSAR, H. E.: *Am. Mineral.* **10**, 123 (1925).
- (9) GOLDSCHMIDT, V. M. (AND OTHERS): *Skrifter Norske Videnskaps. Akad. Oslo.*, 1 *Matemat. natur. Klasse* **1923**, No. 3; **1924**, No. 4, 5; **1925**, No. 5, 7; **1926**, No. 1, 2.
- (10) PIUTTI, A.: *Rend. accad. sci. Napoli* **31**, 72-3 (1925).
- (11) FERNANDES, L. *Atti. accad. Lincei* (6), **2**, 182-6 (1925).
- (12) URBAIN, E. AND G.: *Compt. Rend.* **178**, 265 (1924)
- (13) PRATT, J. H.: *Bull. No. 25, N. C. Geol. and Econ. Survey*, 1916.
- (14) KUNZ, G. F.: *Bull. No. 12, N. C. Geol. and Econ. Survey, History of the Gems Found in North Carolina* (1907), p. 52.
- (15) LIDDELL, D. M.: *Eng. Mining J.* **104**, 153-5 (1917).
- (16) WATSON, T. L., AND HESS, F. L.: *Zirconiferous sandstone near Ashland, Va.*, Univ. of Va. Pub., *Bull. Philos. Soc., Sci. Section*, Vol. I, No. 11, pp. 267-292, July, 1912; *U. S. Geol. Surv. Bull.* 530 P.
- (17) HEVESY, G.: *Kgl. Danske Videnskab. Selskab Math.-fys. Medd.*, **6**, No. 7, 4-149 (1926).

- (18) Bull. U. S. Geol. Survey, No. 74 (1891), p. 49 (The Minerals of North Carolina).
- (19) GORDON, S. G.: The Mineralogy of Pennsylvania, Phil., 1922, p. 90.
- (20) MALLET, J. W.: *Am. J. Sci.* [3], **14**, 397 (1877).
- (21) HIDDEN, W. E.: *Am. J. Sci.* [4], **19**, 432 (1905).
- (22) HIDDEN, W. E., AND MACKINTOSH, J. B.: *Am. J. Sci.* [3], **38**, 1889, 485 (1889).
- (23) LUQUER, L. McL.: *American Geologist* **33**, 17 (1904).
- (24) JAMES, C.: *Private communication*, Feb. 10, 1925.
- (25) HEVESY, G.: *Private communication*, Oct. 31, 1926.
- (26) MEYER, H. C.: Brazilian Zirkite Deposits, Monthly Prices for November, Foote Mineral Co., 1916, pp. 29-31, quoted by W. T. Schaller, Zirconium and Rare Earth Minerals, *Mineral Foote-Notes*, Vol. II, No. 3, March, 1918, p. 2.
- (27) SHEPARD, C. U.: *Am. J. Sci.* [2] **37**, 405 (1864).
HIDDEN, W. E.: *Am. J. Sci.* [3] **38**, 494 (1889).
WILLIAMS, J. F.: *Ibid.*, **40**, 457, 1890 (Eudialyte and eucolite from Magnet Cove, Arkansas).
- (28) FERSMAN, A. E.: *Am. Mineral.* **11**, 289 (1926).
- (29) GORDON, S. G.: *Eng. Mining. J. -Press*, Feb. 6, 1926.
- (30) STEENSTRUP, K. J. V.: *Z. Kryst. Mineral.* **50**, 283-4 (1912).
- (31) JAMES, C.: *Private communication*, Oct. 20, 1926.
- (32) HEVESY, G., AND JANTZEN, V. T.: *J. Chem. Soc. T* **123**, 3218 (1923); *Chem. News*, **128**, 341 (1924).
HEVESY, G.: *Rècherches sur les propriétés du hafnium* (København, 1925).
- (33) International Critical Tables, Vol. 1, p. 114 (1926).
- (34) MEYER, H. C.: Zirconia, its occurrence and application. Paper read before the Ceramic Society, Swansea, Wales, September, 1918. *Mineral Foote-Notes*, March-April, 1919, pp. 2-12.
- (35) VAN ARKEL, A. E., AND DE BOER, J. H.: *Z. anorg. allgem. Chem.* **148**, 345-50 (1925).
- (36) *B. P.* 211, 825; 216, 840; 220, 301

General references

- (37) LANGMUIR AND BASKERVILLE: *Index to the Literature of Zirconium* (1789-1898), Smithsonian Misc. Coll. (1899), No. 1173
- (38) VENABLE, F. P.: *Zirconium and Its Compounds*, Chemical Catalog Co., N. Y. (1922).
- (39) HEVESY, G.: *Kgl. Danske Videnskab. Selskab Math.-fys. Medd.* **6**, No. 7, 4-149, København, A. F. Høst & Søn (1926).
- (40) HEVESY, G.: *Chemical Reviews*, **2**, 1 (1925). (Includes bibliography of 63 references.)
- (41) ROSE, H.: *Das Hafnium* (Braunschweig, 1926).
- (42) DOELTER AND LEITMEIER: *Handbuch der Mineralchemie*, Band III, Zweite Hälfte, p. 1170 (Steinkopff, 1926).
- (43) MELLOR, J. W.: *A Comprehensive Treatise on Inorganic and Theoretical Chemistry*, Longmans & Co. (1927), Volume VII, Chapter XLII, Zirconium. Volume VII, Chapter XLIII, Hafnium.
- (44) HEVESY, GEORG: *Das Element Hafnium*, Berlin, Julius Springer (1927).
- (45) CAHEN, E., AND WOOTTON, W. O.: *The Mineralogy of the Rarer Metals*, Griffin & Co. (1920).

THE RADIATION HYPOTHESIS OF CHEMICAL REACTION

FARRINGTON DANIELS

Laboratory of Physical Chemistry, University of Wisconsin, Madison

Few hypotheses in science have suffered such a rapid rise and fall as the radiation hypothesis. Attracting attention soon after the close of the war, it was eagerly taken up by theoretical chemists because it offered a new point of attack for an important problem which had hitherto defied solution. In contrast with the remarkable successes of thermodynamics in determining how far a chemical reaction can go, nothing but failure had come from attempts to predict how fast it will go. Any hint, however slight, was worth following in this attempt to solve the problem of reaction velocities. The radiation hypothesis seemed particularly important because it suggested a correlation between the fields of chemical kinetics, photochemistry and atomic structure,—between reaction rates, absorption spectra and activated molecules.

The names of Trautz (1), Perrin (2) and W. C. McC. Lewis (3) are associated with the development of the original radiation hypothesis. The best statements of the hypothesis were given at a symposium held in 1922 (4). Already vital objections had been raised and at a similar symposium (5) on photochemistry three years later there was little support for the remnants of the original hypothesis. Several good reviews of the hypothesis are available for work up to 1923 (6).

Experimental evidence for testing the hypothesis was lacking at first because very few reactions exist which do not have complicating phenomena such as molecular collision or catalytic effects, which mask the fundamental process. The experiments which have been devised recently to test the hypothesis have shown the original hypothesis to be untenable, and leave one

which is so difficult as to defy immediate experimental proof and so complicated as to be of doubtful value if proved to be correct.

The experimental tests of the radiation hypothesis have been concerned chiefly with unimolecular reactions, because such reactions were supposed to be independent of molecular collision; and with infra-red spectroscopy, because infra-red radiation is the most abundant at ordinary temperatures, and because quantum calculations indicated that it was the effective radiation for most reactions which proceed with measureable velocity.

THE STATEMENT OF THE HYPOTHESIS

In spite of the apparent failure of the original hypothesis, certain things of importance have been brought out by its formulation. It is necessary first to state the hypothesis in its broadest terms, and the various deductions which have been associated with it.

There are four main postulates connected with the original radiation hypothesis of chemical reaction.

First postulate. Molecules must be put into an active state before they can react chemically, and in ordinary thermal reactions this activation is brought about by radiation emitted by the walls of the containing vessel.

Second postulate. The energy per gram molecule required to activate the molecules, E , (called the critical increment) may be calculated from the temperature coefficient of the reaction rate, k , by the equation

$$\frac{d \log_e k}{dT} = \frac{E}{RT^2}$$

Third postulate. Assuming that each molecule is activated by one quantum, $h\nu$, the energy absorbed in activation is given by the expression

$$E = N h\nu.$$

where N is the Avogadro number (6×10^{23}), h is Planck's constant (6.57×10^{-27}) erg—seconds and ν is the frequency of light. By this equation and that of the second postulate the frequency

of the activating radiation can be calculated from the temperature coefficient of the reaction rate. Conversely, too, the temperature coefficient of the reaction rate can be calculated from the frequency of the light which brings about the reaction.

Fourth postulate. The relation between chemical reaction and radiation is reversible, and radiation of the activating frequency is emitted when the reaction is reversed.

UNIMOLECULAR REACTIONS

The first postulate was suggested in 1913 by Perrin (7). He pointed out that unimolecular reactions are independent of collision, for if one doubles the volume of a gas which is decomposing unimolecularly the number of collisions will be only one-fourth as great, and yet the same number of molecules decompose per second as before. This independence of concentration is a requirement of a unimolecular reaction. Perrin reasoned that since the molecules in such a reaction are independent of collisions, they cannot get their energy of activation from collisions and they must get it from radiation emitted by the walls of the containing vessel. Obviously such a conclusion is valid only in case collision and radiation are the only possibilities and in case the exclusion of collisions is fully established.

It was not long before critical examinations and new experiments showed that true unimolecular reactions are extremely rare. One by one, the supposedly unimolecular reactions were shown to be more complicated reactions, for they were found to be catalyzed by the walls of the containing vessel or by traces of impurities. It was pointed out by Rice (8) that if there are no true unimolecular reactions there is no need to invent a theory to account for them.

For some time the decomposition of nitrogen pentoxide (9) remained the "sole survivor of the group of truly unimolecular reactions" and for a while even this reaction was in doubt. Later experiments, (10), (11), (12), however, have confirmed the true unimolecular character of the reaction. Very recently this reaction has been critically studied to determine if it involves catalysis in any way either from dust particles of phosphorus

pentoxide, or moisture or other material. It has again been pronounced a truly unimolecular reaction, uncatalyzed and homogeneous (13).

The existence of unimolecular reactions has been supported by the study of several other reactions as follows: The decomposition of sulfuryl chloride (14), acetone (15), propionic aldehyde (16), di-ethyl ether (17), di-methyl ether (18), azo methane (19a, b), and by the gas phase racemization of pinene (20). All of these reactions have been shown to be unimolecular, homogeneous, gas phase reactions, although in some cases the specific reaction rate falls off at low pressures.

The fact is now fully established that unimolecular reactions do exist, and they must be explained. In the next section the collision hypotheses advanced to account for unimolecular reactions are reviewed briefly.

COLLISION HYPOTHESES

Perrin believed that the collision theory was ruled out by the fact that the reaction rate is independent of the collision rate.

Lindemann (21) suggested that there may be a time lag between activation and decomposition which would make the reaction rate appear to be independent of the number of collisions. Under these conditions the activated molecule can have its energy taken away again by collisions before it has had time to decompose. The rate of decomposition, however, would still be proportional to the concentration and the unimolecular constant would be satisfied, except at low concentrations when the time between decompositions would be of the same order as the time between collisions. It has been generally supposed, however, that decomposition is synonymous with activation in a true unimolecular reaction. A comparison of the racemization of pinene with the decomposition of nitrogen pentoxide (20), supported the view that the time between activation and decomposition is negligible at least in these two reactions.

Rodebush (22) suggested a statistical theory for unimolecular reactions according to which the dissociation of a molecule is caused by a series of collisions, fortuitously timed so as to ac-

celerate the oscillation of some part of the molecule to the breaking point.

According to these hypotheses a break-down of the unimolecular constant is to be expected at very low pressures, so low that the period between collisions is greater than the average calculated time between the decomposition of the molecules.

Attempts were soon made to ascertain if the unimolecular rate constant remains unchanged at low pressures in accordance with these ideas. The experimental errors are seriously magnified in this region. White and Tolman (10) and Hunt and Daniels (11) found no change in the specific decomposition rate of nitrogen pentoxide. Hirst and Rideal (23) reported an *increase* of nearly four fold at very low pressures but this research was quickly checked in other laboratories and the results were not confirmed. Hibben (24) and Smith and Loomis (25) found that the specific rate of decomposition at low pressures checks exactly with that of Daniels and Johnston (9) at higher pressures. Hibben carried the measurements down to 0.002 mm. without observing any change and Smith and Loomis concluded that the results of Hirst and Rideal were complicated by the absorption of N_2O_5 by solid N_2O_4 .

Although nitrogen pentoxide maintains its specific decomposition rate over a very wide range of pressures, Hinshelwood found that the specific decomposition rate of the unimolecular reactions which he has reported (15), (16), (17), (18) decreases at lower pressure, and in most cases becomes di-molecular. Rampsberger (19b) found that the specific decomposition rate of azo-methane decreases also at lower pressures. These changes in the specific decomposition rate occur, however, at pressures of a few millimeters of mercury where each molecule suffers a very great number of collisions before it undergoes decomposition, and no one has succeeded yet in studying the reaction rates at pressures so low as to make the collision frequency as slow as the expected frequency of decomposition.

J. J. Thomson (26) suggested a mechanism, based on collisions, which demanded that the presence of an inert gas should increase the decomposition rate, but this prediction has not been supported

by the experiments with nitrogen pentoxide (10), (11), (12). Hinshelwood (18), found that the decomposition of propionic aldehyde, di-methyl ether, and di-ethyl ether is accelerated by hydrogen but the reaction appears to be connected with some specific property of hydrogen.

The fact that the specific reaction rate of a unimolecular reaction is independent of concentration, constituted only a questionable argument against the collision hypothesis. A stronger argument against the hypothesis was based on the fact that ordinary collisions cannot supply energy rapidly enough to bring about the decomposition.

These calculations have been made by Christiansen and Kramers (27) by Lewis and Smith (28) and by Tolman (29) and all are agreed that the most natural source of energy, the collision between two ordinary molecules is insufficient to cause decomposition in a unimolecular reaction. The numbers of collisions can be calculated from the kinetic theory of gases with reasonable estimates of the diameters of the molecules. The number of collisions giving *sufficient* energy for the reaction is of the order of one-ten thousandth of the number required to check with the experimentally determined constant.

INTERNAL ENERGY OF MOLECULES

Failing to explain unimolecular reactions by ordinary collisions, workers in this field have now turned to other sources of energy. Lewis and Smith (28) suggested that the large heat capacity of the molecule of nitrogen pentoxide may contribute some energy from previous collisions.

Hinshelwood (30) pointed out that as far as our evidence goes now it is only the complicated molecules with many degrees of freedom which have been found to decompose according to a unimolecular reaction, and this observation led him to the suggestion that the internal energy of the molecule can contribute to the energy of decomposition. Since it is necessary to combine a collision of high energy with an internal arrangement of high energy (or low stability) in order to bring about decomposition, the effect of collisions is largely masked. Hinshelwood suggested

also that the unimolecular decomposition may demand the liberation of free atoms, and this process requires the absorption of a large amount of energy. Only the complicated molecules with many degrees of freedom can supply sufficient energy.

Several attempts have been made to formulate quantitative expressions for the unimolecular reaction based on more elaborate collision hypotheses. In general these elaborated hypotheses make use of the internal energy of the molecule and several degrees of freedom. Roy (31) emphasized the similarity between photo-electric phenomena and photo-chemical reactions and proposed an equation for a unimolecular reaction which involved the molecular weight, the heat of dissociation, the cross section of the reacting molecule and an arbitrary constant connected with the number of valence bands.

Christiansen (32) criticized some of the deductions of Lewis and Smith (28) but emphasized the need for considering several degrees of freedom.

J. J. Thomson (33) proposed a theory for unimolecular reactions based on the assumption that the law of the conservation of energy applies as a statistical average to a group of molecules, but that it does not apply to isolated molecules if the forces inside the molecule are discontinuous. This theory led to an expression which gives an increase in the specific reaction rate at low pressures and checked the experiments of Hirst and Rideal (23) but these experiments were not confirmed (24), (25).

Fowler and Rideal (34) in a comprehensive treatment discussed the ways in which molecules can be activated by collision with sufficient rapidity. They assumed that the energies of two colliding molecules can be used. Assuming 15 degrees of freedom in the nitrogen pentoxide molecule they showed that it is possible to account for the observed decomposition rate and in fact their calculations give a rate twenty times too large. They pointed out that the target area of a molecule for de-activation may not be the same as the target area for activation.

Hinshelwood (35) and Hinshelwood and Askey (36) found that the decompositions of propionic aldehyde, di-methyl ether and di-ethyl ether were unimolecular at higher pressures but at

pressures of a few millimeters they became dimolecular. The addition of hydrogen caused the unimolecular decomposition rate to be maintained down to the low pressures but helium and nitrogen did not have this effect. The authors give the name "quasi-unimolecular" to reactions which give unimolecular constants at higher pressures maintained, apparently, by collisions. They give a correction term for calculating the energy of activation, based on the number of degrees of freedom.

Rice and Rampsberger (37) suggested two theories based on the falling off of the unimolecular constant of propionic aldehyde at low pressures and the utilization of the internal degrees of freedom of complex molecules. The first theory assumes a minimum value of energy necessary for reaction and the second theory a minimum value *in a particular degree of freedom*, and each one fitted the experimental data about equally well.

Final judgment of these various hypotheses must await further experimental tests but at present it seems quite probable that unimolecular reactions can be explained satisfactorily on the basis of collisions combined with several degrees of freedom inside the molecule. More unimolecular reactions are needed as well as specific heat measurements and data from which the diameters of the molecules can be calculated. In the simpler molecules there is no time lag between activation and reaction but in the more complex molecules with several degrees of freedom the activated molecule may suffer many de-activating collisions before it has time to react. In the former case a dimolecular reaction, dependent on collision, is the result. In the latter case the time lag obliterates any apparent relation between collision frequency and rate of reaction although the collisions still provide the necessary energy. The number of molecules which reach the state, necessary for decomposition, is proportional to the total number present and thus the mathematical requirements of the unimolecular reaction are met.

Certainly these collision hypotheses are sufficiently promising so that it is not necessary to bring in radiation as an activating mechanism for these reactions.

CHAIN REACTIONS

In the preceding discussion only collisions between ordinary molecules, have been considered. Another promising source of energy which might be used to bring about the unimolecular reactions, lies in collisions with activated molecules, the so called collisions of the second class. When a molecule is in an activated condition with a displaced electron, it is able to pass over its extra energy to other molecules with which it collides. In this way energy produced in the chemical reaction can be used over again for the activation of the molecules. In other words it may not be necessary to account for the whole energy of activation but only for the heat of reaction, which in most cases is much less.

Christiansen and Kramers (27) have developed this idea to the fullest extent. They propose specific reaction chains whereby the energy liberated by the decomposition of the activated molecule is passed on to new, reacting molecules. In this way a steady state is reached with the number decomposing, proportional to the total number of molecules.

Several objections have been raised to this mechanism. First, the decomposition of a complex molecule is usually associated with the absorption of energy rather than with the evolution of energy. In the case of the decomposition of nitrogen pentoxide, at least, the whole reaction appears to be endothermic.

Second, the presence of inert molecules might be expected to absorb the energy of activated molecules and prevent its return to the reacting molecules, thus slowing down the reaction. However, in the case of nitrogen pentoxide the reaction rate is not appreciably affected by a million fold excess of nitrogen molecules (11). Even when surrounded by an inert solvent the molecules of nitrogen pentoxide decompose at nearly the same rate (38). These objections may be somewhat discounted by the fact that the action of activated molecules is likely to be quite specific, as shown by experiments on the quenching of fluorescence.

A third argument against the chain reaction is raised by Hirst

and Rideal (23), who studied the decomposition rate of nitrogen pentoxide at very low pressures. The pressure was so low that each molecular chain, if it existed, would be subjected to impacts on the wall some forty times, and one impact should be sufficient to break it. In spite of these arguments it seems reasonable that some of the energy of activation can be contributed by the products of the reaction.

Norrish (39) found that nitrogen dioxide is decomposed photochemically into nitric oxide and oxygen and Busse and Daniels (40) found that nitric oxide reacts immediately with nitrogen pentoxide. These facts suggest that in the thermal reaction half of the nitrogen pentoxide molecules decompose spontaneously according to a slow unimolecular reaction but that for each molecule decomposed a second molecule of nitrogen pentoxide is decomposed by the reaction products. In this way it is possible to avoid an equation for the reaction which involves the unlikely liberation of atomic oxygen, and yet permits the equation to be written with a single molecule of nitrogen pentoxide decomposing in accordance with the experimentally established unimolecular decomposition (40). The arguments for this mechanism seem sufficiently strong to emphasize the hypothesis that some of the energy for decomposition of the nitrogen pentoxide may be obtained by reaction between the decomposition product (nitric oxide) and half of the original nitrogen pentoxide, as proposed originally by Christiansen and Kramers (27). It is probably significant that most of the unimolecular reactions fall off at low pressures except nitrogen pentoxide, and in this decomposition there is this evidence for a secondary reaction. It will be necessary to study further the thermochemistry of the nitrogen oxides before settling the question definitely.

ACTIVATED MOLECULES

At the present writing the existence of unimolecular reactions does not seem to demand activation by radiation, but it does appear to demand some kind of an activated molecule. The chief argument for the existence of activated molecules is the fact of slow reactions, for reactions should take place instantly if all molecules are equally active.

Arrhenius (41) first postulated active and passive molecules to account for the fact that a slow chemical reaction increases 20 or 30 per cent per degree rise in temperature. A shift in the equilibrium from passive to active molecules could then explain the large temperature coefficient of chemical reaction.

There has been no direct experimental proof of the existence of activated molecules in ordinary thermal chemical reactions except perhaps the spectral evidence of decomposing ozone pointed out by Wulf (42) and the influence of water vapor in quenching the infra-red emission spectrum of explosions discussed by David (43). The argument of Arrhenius based on the large temperature coefficient of chemical reactions can probably be explained with the help of statistical mechanics without the necessity of considering radiation.

On the other hand the conception of activated molecules has received a tremendous impetus through the development of the Bohr theory, for it is now easy to visualize an activated molecule, as a molecule with a displaced electronic orbit. It is probable that activated molecules may be considered best in a statistical sense with transitory displaced orbits, rather than new and stable molecules in the sense of structural organic chemistry.

THE CRITICAL INCREMENT

According to the second postulate of the radiation hypothesis it is possible to calculate from the temperature coefficient of the reaction rate, k the energy, E_c , which is required to activate the molecule and prepare it for reaction. This energy is called the critical increment, but there is some difference of opinion as to its exact definition. Most workers seem to favor the definition that it is concerned with average values, and the following definition is, perhaps, the most comprehensive. According to Tolman (44)

$$\frac{d \ln k}{dT} = \frac{E_{(\text{activated})} - E_{(\text{average})}}{RT^2} = \frac{E_c}{RT^2}$$

where E is a term which contains the energy of the molecules and also the energy of the "modes of vibration" entering the

reaction. R is the molar gas constant and T is the absolute temperature.

This equation seems to rest on a firm theoretical foundation. It was derived on the basis of statistical mechanics by Marcelin, (45), by Rice (46) and by Tolman (44). It is based on Maxwell's distribution law, and Hamilton's equations of motion. In Tolman's derivation the influence of radiation is considered.

Only a few experimental data are available for testing this equation directly. These tests, though, as far as they go seem fairly satisfactory.

Perhaps the most striking support for its validity involves the equation

$$k = S e^{\frac{-E}{RT}}.$$

In this equation k is the specific reaction rate, E is the critical increment, R is the gas constant and T is the absolute temperature. $e^{\frac{-E}{RT}}$ represents the fraction of the molecules which have sufficient energy to become activated. E is calculated by integrating the expression given earlier in this section. Experimentally, then, E may be determined from a knowledge of k at two different temperatures. Now in dimolecular reactions, S turns out to be numerically equal to the number of collisions as calculated from the classical formulas of the kinetic theory of gases and this equality could hardly be possible unless the calculation of E is correct. Also the ratio, $\frac{E}{T}$, is constant for widely different reactions, and this constancy could not be expected unless there is something significant about the calculation of E . Even in unimolecular reactions where the significance of S is not now apparent, $\frac{E}{T}$ is constant. For example (47), nitrogen pentoxide, at 55°C. gives $k = 0.00015$ and $E = 24,700$, and acetone at 562°C. gives $k = 0.00015$ and $E = 68,500$. For nitrogen pentoxide $\frac{E}{RT} = 38$ and for acetone $\frac{E}{RT} = 41$, when the temperatures are chosen so as to give equal reaction rates, i.e. the same value of k .

Another check on the calculation of E is possible through the relation between the equilibrium constant, K , and the specific reaction rates, k_1 and k_2 , of the two opposing reactions, as given by the equation

$$K = \frac{k_1}{k_2}$$

According to the familiar van't Hoff equation, the heat of reaction, Q , and the equilibrium constant, K , are related as follows,

$$\frac{d \ln K}{dT} = \frac{Q}{RT^2}$$

combining this equation with the equation involving the critical increment it can be shown that $Q = E_1 - E_2$. In other words, the heat of reaction, calorimetrically measured, is equal to the difference between the energy of activation of the direct reaction and the reverse reaction. Indirect confirmations of this relation have been pointed out by Dushman (48). It would be very desirable to check this relation in significant reactions in which the temperature coefficient of the two opposing reactions, and the heat of reaction are all determined experimentally under the same conditions.

A third method for testing the validity of calculations of the critical increment is found in quantitative photochemical investigations. The photochemical energy required to cause the reaction of a gram molecule of material should check with the energy of the critical increment if secondary reactions are eliminated.¹

THE QUANTUM HYPOTHESIS

According to the third postulate of the radiation hypothesis, given earlier,

$$E = N h \nu.$$

¹ The results of such an experiment on nitrogen pentoxide will be reported in a later publication.

This equation involves the quantum theory and it must stand or fall with this theory. It will be noted that this equation is identical with the Einstein law of photochemistry, which seems to be gaining slightly in significance as reactions free from complications are found and experimental technique is improved.

This postulate is the one which has been most often discussed in connection with the radiation hypothesis and it was the one which first discredited the hypothesis.

The temperature coefficient. It was pointed out first that this equation leads to the requirement that a single wave length of light is effective in bringing about the activation of the molecules. Perrin (2) next showed that the temperature coefficient of the emission of monochromatic light by a black body is of the same form as the temperature coefficient of the rate of a chemical reaction. In both cases a straight line results when the logarithm of the property is plotted against the reciprocal of the absolute temperature. Perrin used this fact as an argument that reaction rate depends on absorption of radiation from the walls of the containing vessel, but Langmuir (49) showed that there is nothing significant about it, because both are probability phenomena and both follow an exponential equation.

According to this relation E should be independent of temperature. In some cases, E does change slightly with temperature but Tolman (44) used this fact as an argument for several frequencies rather than the single frequency, ν .

Originally the large difference between the temperature coefficient of reaction rate and collision frequency was considered an argument against the collision theory and therefore an argument in favor of the radiation hypothesis. However, this argument fails when it is realized that only one collision in many millions can give sufficient energy for the activation, and this frequency of unusually violent collisions is, by Maxwell's distribution law, an exponential function.

Insufficient radiation density. Against this quantum equation Langmuir (49) also raised the objection that the monochromatic radiation density is by no means sufficient to supply the energy demanded by the critical increment. Similar calculations have

been made also by several others (27), (28), (29) and there seems to be no doubt that the walls of the containing vessel fall short of giving the necessary radiation by at least a million fold, when calculations are based on the black body radiation laws.

Several ways out of this difficulty have been proposed, but they are not at present convincing.

Tolman (50) proposed an elaborated radiation hypothesis in which the reacting material is activated by several frequencies instead of by one. In this way more radiant energy from the walls can be used.

It must be confessed that we know very little about the radiation *inside* a "hohraum," for all our experience is based on phenomena in which the radiation comes outside. Under the conditions of the chemical reaction the inside of the chamber is practically absolute zero as regards the radiation of the absorbable frequencies and yet it is up to full radiation temperature as regards the remaining spectrum. Under these conditions it is, perhaps, unsafe to use the classical laws of black body radiation in the ordinary way.

In view of our ignorance of radiation inside a chemically reacting system, Lewis and McKeown (51) have suggested that it is possible to draw off energy of a single frequency at very much greater rates than calculated by the ordinary radiation laws. In fact they suggested that the energy for all the other frequencies can be converted instantly into the one which is being used up and that the total radiation of the black body is available for chemical reactions.

In discussing the properties of radiation in a chemically reacting system it is necessary to distinguish clearly between a thermal and a photochemical reaction. It has been pointed out (51), (52), (31) that in a photochemical reaction the radiant energy comes from an outside source and is not in equilibrium with the system. In a thermal reaction, however, the radiation is in thermal equilibrium with the walls of the containing vessel.

The difficulty of finding sufficient radiation density may be removed in a third way, by considering that part of the energy of activation is provided by collisions, or by secondary reactions.

In this way it is conceivable that the radiation would have to supply energy corresponding to only a fraction of the critical increment. If the heat of the reaction, i.e. the difference between the critical increments of the products and the reacting material determines the energy requirement, this difficulty of insufficient radiation density can be minimized or nullified. The help offered by this consideration can be used for the collision hypothesis as well as for the radiation hypothesis.

Absorption spectra. A strong objection to the quantum postulate of the simple radiation theory is found in the evidence of absorption spectra. In many cases there is no absorption at the frequency, ν , demanded by the equation $E = N h\nu$. The most striking example is the case of nitrogen pentoxide which is a true unimolecular reaction with a critical increment of 24,700 calories well confirmed by different workers. Setting $E = N h\nu$, the activating frequency ν is in the near infra-red at 1.16μ but no absorption of light can be found in this region (52). Other examples of a similar nature can be cited.

The difficulty can be met by Tolman's elaborated radiation hypothesis (50) in which several frequencies are involved.

This quantum relation can be tested also in another way. Nitrogen pentoxide decomposes in different solvents at rates which are only slightly different than those in the gas phase. The critical increment varies slightly for different solvents and should lead to a corresponding shift in the lines of the absorption spectrum. No such shift was found (53) although the calculated shift should have been large enough to measure with certainty.

It is possible that some of the failures to check the calculated frequency, with the absorption spectrum, may be due to a confusion of several critical increments. If the experimentally determined reaction rates on which the calculation of E and ν depend, consist of several different reaction rates, the calculated frequency ν is not significant. If side reactions and consecutive reactions cannot be separated from the main reaction there is little hope of calculating the significant critical increment. This situation has been emphasized particularly by Garner (54).

Photochemical inactivity of infra-red radiation. The greatest

objection to the simple radiation hypothesis lies in the fact that the calculated frequency of radiation is not able to bring about the reaction, photochemically. If this radiation, brought in from an outside source of great intensity, is inactive, then the radiation from the containing walls at a much lower temperature should not be effective.

The nitrogen pentoxide reaction again furnishes the best experimental data now available. As pointed out in the preceding section, the simple radiation hypothesis, demands that light of 1.16μ should decompose nitrogen pentoxide. Experimental facts (52), (58) show that no decomposition is effected even by very intense light of this frequency.

Although there is no absorption band at 1.16μ there are a few absorption bands in the infra-red and one of these falls at $3 \times 1.16\mu$ and another at $5 \times 1.16\mu$ (55), (53). It was thought that this fact might be significant (56) and that radiation of these wave lengths would be effective. It was shown, however, that these radiations are also inactive (57), (58). These multiples of the calculated wave length then appear as coincidences. Intense radiation from a Nernst glower was passed through a transparent bulb containing nitrogen pentoxide and it was shown that at least 99 per cent of the energy actually absorbed was without effect on the decomposition of the nitrogen pentoxide (58).

In the same research it was found that intense infra-red radiation does not decompose carbon dioxide nor hydrochloric acid gas and does not bring about the reaction between alcohol vapor and oxygen. If the radiant energy actually absorbed is divided by a very large, assumed critical increment, the quotient corresponds to thousands of micromoles, whereas the tests were sensitive to one micromole and less.

These experiments were concerned with the near infra-red from 2μ to 7μ , and Lewis and Mayer (73) have found recently that far infra-red radiation as well as the near infra-red is ineffective in decomposing nitrogen pentoxide. Hibben (24) found that infra-red radiation is ineffective also in decomposing ozone and nitric oxide. Although the cases now available for test are too few for extensive generalization, there is up to the present

time no established case of a chemical reaction brought about by infra-red radiation, except possibly the influence of infra-red radiation on explosions (43). This situation is to be expected for,

According to present theories absorption bands in the near infra-red are due to the displacement of atoms within the molecule, while visible and ultra-violet absorption lines are due to the displacement of electrons. The latter process results frequently in chemical decomposition, but there is no evidence that the former does. It is suggested that the two processes are fundamentally different in mechanism and that atoms may be loosened in a molecule by displacement of the electrons which hold them together, but not by direct displacement of the atoms (58).

Apparently the infra-red radiation does not contain sufficiently large enough quanta, i.e. high frequency to cause activation.

When a reaction is already taking place thermally, infra-red radiation can cause slight acceleration of the reaction rate but this effect then becomes indistinguishable from an ordinary thermal effect. This phenomenon was noted in the experiments with nitrogen pentoxide at room temperatures (58).

The apparent inactivity of infra-red radiation is a serious blow to the radiation hypothesis for it was to the infra-red radiation from the walls of the containing vessel that most of the importance was attached.

THE ELABORATED RADIATION HYPOTHESIS

Although the simple radiation hypothesis has been proved untenable, it is still possible that an elaborated radiation hypothesis may survive.

Such an elaborated hypothesis has been nicely developed by Tolman (59) who discusses the following possibilities.

Simultaneous absorption of several quanta. If several quanta are involved in the passage from the normal to the activated state, the frequency of the activating light cannot be calculated by setting the critical increment equal to $Nh\nu$, because several frequencies are involved. Perrin has advanced this explanation to explain some of the failures of the simple radiation hypothesis.

This suggestion seems unlikely because coupled absorption of this type is not to be expected from the quantum theory and it is not supported by spectroscopic evidence. Moreover the chance for simultaneous collisions with two or more quanta is much less than the chance for collision with one.

Successive absorption of quanta. It is quite possible that the molecule does not pass from a normal to an activated molecule in a single change, but rather that it passes through a series of intermediate stages. Several smaller quanta of different frequencies would then be required and the discrepancy between calculated and observed absorption bands could be explained. This mechanism, however, is of no help in getting around the difficulty of insufficient radiation density.

Simultaneous absorption and emission of quanta. If quanta are emitted simultaneously with the absorption of quanta the radiation in a closed space may suffice to give reaction rates which agree with those experimentally determined, but there is no evidence now to show that this process is important.

Absorption of a range of frequencies. Instead of a single activated state, there may be a whole range of activated states of continuously increasing energy content. Such an increase in the number of activated states would help greatly in the problem of obtaining sufficient energy to account for the observed reaction rate, but it is not possible now to test the hypothesis quantitatively. This modification of the simple radiation hypothesis appears to be the least hopeless.

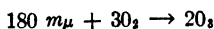
Another way out of the difficulty may possibly be found in the suggestion of Polanyi (60) that a coupling through radiation exists between the molecules decomposing and those being activated.

THE EMISSION OF LIGHT

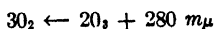
According to the fourth postulate of the radiation hypothesis, radiation of the activating frequency is emitted when the reaction is reversed. The idea can best be illustrated with a specific case.

It has been found experimentally that light of $180\text{ m}\mu$ will

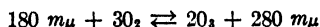
convert oxygen into ozone and that light of $280\ m\mu$ will convert ozone into oxygen, according to the reactions,



and



These two equations may be considered together and written as an ordinary balanced reaction



The reversibility of such a process seems perfectly natural. If the complete reversibility is correct, it follows that light of $280\ m\mu$ is emitted when ozone is formed, and light of $180\ m\mu$ is emitted when ozone is decomposed.

In general



where ν and ν_1 , are different frequencies, and A and A' are the reactants and the products. Similar equations would apply also to polymolecular reactions.

Perrin (2) and Lewis (3) have pointed out how simply this hypothesis will explain the phenomena of chemiluminescence, fluorescence and phosphorescence. The emission of light is, in fact, a necessary requirement of chemical reaction and fluorescence is simply a special case in which the chemical reaction is brought about by radiation from an outside source. Phosphorescence is similar to fluorescence except that a time lag is involved in the chemical reaction. There is something very attractive about this fourth postulate of the radiation hypothesis.

According to the original radiation hypothesis all chemical action is accompanied by the absorption of radiation and all chemical reaction involves the emission of radiation. Indeed, it is strange that chemiluminescent reactions are not more common. If the absorption of radiation, however, is not a necessary precursor for chemical reaction, then the simple radiation hypothesis falls and the emission of light is not a necessary accompaniment of chemical reaction.

It must be pointed out that the experimental tests for this emission of light are difficult. The light intensity is very low in slow reactions and only the light emitted near the outside of the reacting mass can be detected. Most of the light is emitted in the interior, where it is absorbed by the molecules of the products before it can get out. Possibly the quantity of light reaching the outside of the reacting system is too small to be detected except in special cases where the absorption coefficients happen to be favorable.

Although several examples of chemiluminescence are known in no case has the frequency of the emitted light been connected quantitatively with the radiation hypothesis.

It is worth while to search for new examples of chemiluminescence in simple reactions where the critical increment can be determined. Also, a study of the temperature coefficients of rates of luminescent reactions may give information concerning the critical increment of luminescent reactions, which are now known. Two tests of the quantum equation are possible as shown by the following equations

$$\text{Frequency of emitted light, } \nu, = \frac{E}{Nh}$$

and

$$\text{Heat of reaction, } Q, = E_1 - E_2 = Nh(\nu_1 - \nu_2)$$

According to the latter equation the difference in the frequency of the emitted light and that of the chemically activating light multiplied by Nh , is equal to the heat of the reaction. The latter test avoids the necessity of determining the critical increments. Direct experimental tests of these equations are lacking at present.

Perrin (61) has studied the fluorescence of dyes, particularly methylene blue in glycerin in an attempt to check relations predicted by the radiation hypothesis. In general, there seems to be some indirect evidence supporting the radiation hypothesis.

Some experiments have seemed to confirm the predictions of the radiation hypothesis but they have been later disqualified. A

case in point is concerned with the absorption spectrum of anthracene and the prediction of certain quantum relations (62), but the absorption bands on which the predictions were based were later shown to be due to impurities (63).

Great care must be exercised in experiments designed to test photographically the emission of light in chemical reactions, for photographic plates are very sensitive to organic vapors and other substances. For example it was reported by Kugelmass and McQuarrie (64) that ultra-violet light is emitted when cod liver oil is oxidized and that oxygen is liberated when the oil is radiated with ultraviolet-light. These experiments were of interest in the study of the cure of rickets and they appeared to constitute an excellent example of the fourth postulate of the radiation hypothesis. The results were not checked (66) however, and the first conclusions were withdrawn (65).

The question to be decided by future experiments is whether radiation is usually emitted in chemical reactions but in such a way that it cannot be easily detected; or whether the energy evolved in a chemical reaction goes immediately into kinetic energy of the molecules.

EXPLOSIONS

Important facts bearing on the question just raised have been discovered in the study of explosions. Light is emitted in flames and explosions and it has been known for some time that a considerable quantity of infra-red radiation is emitted. These are violent, rapid reactions and here if anywhere the emission of radiation should be sufficiently great to detect. The fact that radiation is emitted in these reactions, constitutes a support for the radiation hypothesis.

The question arises as to whether this emission of radiation is really chemiluminescence or whether it is thermal emission. In the first place we know very little concerning the emission of radiation from a heated gas in the absence of chemical reactions or electrical ionization. More experimental data in this field would be very valuable.

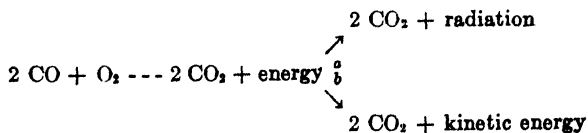
The following considerations show that some of the radiation at least is due to chemiluminescence:

1. David (67) found that the emission of radiation is greatest before the maximum temperature of the explosion is reached. If all the radiation is of thermal origin the greatest emission should come at the highest temperature. The most rapid chemical reaction occurs just before the maximum temperature is reached and it is here that the emission of radiation is the greatest.

2. Haslam, Lovell and Huneman (68) discovered that the preheating of the gases decreases the emission of radiation. If the radiation is due to temperature, the preheating should have increased the radiation. If the radiation is of chemical origin the preheating might decrease the emission by causing a more rapid attainment of equilibrium with a quicker degradation of chemical energy into kinetic energy.

Bone and Webster (69), found that water vapor decreases the radiation by catalyzing the reaction. Garner and Johnson (70) found that the drying of the gases decreased the temperature of the explosion and increased the emission of radiation. The possibility that water vapor acted as a screen for the infra-red radiation was carefully considered, and the conclusion was that less radiation is emitted when the reaction proceeds rapidly and allows the energy of the reaction to be dissipated as heat in molecular collisions. This view is further strengthened by the fact that negative catalysts such as carbon tetrachloride and nitrogen dioxide which decrease the speed of the explosion, increase the emission of infra-red radiation.

The following equation (70) seems to account satisfactorily for the data, and to strengthen the view that radiation is emitted in chemical reactions as predicted by the radiation hypothesis, if the conditions are right



The products of the oxidation of carbon monoxide contain extra

energy from the chemical reaction and this energy may be released in two ways, as indicated by *a* and *b*. When the reaction is rapid most of the energy is converted into increased kinetic energy by collisions between these activated molecules and other molecules. When the reaction is slower and the energy is not dissipated immediately in collisions, there is an opportunity for the activated molecules to lose their energy by radiation as indicated by the reaction, *a*. In ordinary, slow, thermal reactions the reaction, *b*, must predominate greatly over the reaction, *a*.

David (43) found that infra-red radiation accelerates the rate of combustion of inflammable gases if the radiation is absorbed. Later experiments (71), however, indicated that the phenomenon is complicated by the fact that the effect was not noted in the absence of nitrogen, and that in some way the reaction involves the oxides of nitrogen.

Certainly radiation and particularly infra-red radiation is emitted in all explosions and this fact lends support to the fourth postulate of the radiation hypothesis.

RADIATION AS A FACTOR

The following hypothetical experiment (72) proves that radiation must be a factor in some chemical reactions, although probably a small factor.

A quantity of ozone is being decomposed photochemically by black-body radiation emitted from a heated enclosure. The ozone is then placed inside the heated enclosure. Certainly the decomposition caused by the radiation will now be just as rapid as before, and probably a great deal more rapid. The fact that the ozone is being decomposed with extreme rapidity by the increased kinetic energy of the heated molecules does not nullify the fact that radiation is still causing some ozone molecules to decompose and that radiation is a factor in the thermal decomposition.

Direct experiments (72) have shown, however, that radiation from the walls of the containing vessel is not the important cause of activation in a thermal reaction. A stream of heated air was

mixed with a stream of nitrogen pentoxide gas in an open room and although the walls of the containing vessel (the room) were at too low a temperature to bring about rapid decomposition, the nitrogen pentoxide decomposed almost as soon as it came in contact with the heated molecules of air. Molecular collision rather than radiation was therefore responsible for this reaction.

Lewis and Mayer (73) found that pinene vapor is not racemized by passing it through the center of a heated tube in a uni-directional stream. This experiment constitutes another decisive argument against the radiation hypothesis because the radiation in the tube corresponded to a temperature sufficiently high to racemize the pinene immediately.

CONCLUSIONS

In conclusion it may be stated that the existence of unimolecular reactions has been fully established but that there is, as yet, no completely satisfactory explanation for their existence. According to different views expressed in the literature, unimolecular reactions cannot be caused by collisions, they cannot be caused by radiation and they cannot be caused by anything but collisions or radiation. Not one of these views can be taken as final.

The hypothesis that radiation is the important factor underlying chemical reaction was originally based on the existence of unimolecular reactions and the difficulty in explaining them on the basis of collision. This argument is not valid now, since several possible hypotheses based on collisions have been brought out by the challenge of the unimolecular reaction and the successes of the past few months in this direction suggest that still more satisfactory theories will be forthcoming. At present the most promising explanation lies in the collision hypotheses and the utilization of the internal energy of the molecule and many degrees of freedom.

The argument for the radiation hypothesis which was based on the large temperature coefficient of chemical reactions is not valid.

The calculation of the critical increment and the conception of

activated molecules seem to be satisfactory and fruitful at the present time. Neither has been disproved experimentally, but on the other hand the experimental evidence in favor of them is meager.

The simple radiation hypothesis involving the quantum theory has been disproved by a number of experiments. These experiments involve (a) the insufficiency of radiation density in a thermal reaction; (b) the discrepancy between calculated and observed absorption bands; (c) the chemical inactivity of infra-red radiation; (d) experiments with cold walls and heated gas, and experiments with cold gas and heated walls.

Objections to the simple radiation hypothesis may be met in several ways, although it must be emphasized that there is now no experimental backing for them.

1. An elaborated radiation hypothesis using many frequencies is possible.

2. Many objections may be removed by admitting complete ignorance of black body radiation laws when applied to chemically reacting systems.

3. The discrepancies may be attributed to complicating intermediate steps, the experimentally determined reaction rate depending on the summation of several reactions.

4. The energy requirements may be much reduced if the products can return part of their energy to the reacting materials.

The reversibility of the relation between chemical reaction and radiation has not been explored very far, and for this reason the field is particularly attractive. Here again, the majority of the evidence does not support the radiation hypothesis, but it cannot be said that this postulate of the radiation hypothesis has yet been disproved. There is something particularly appealing in this postulate, because of its "reasonableness" and because of the ease with which it explains the phenomena of chemiluminescence and fluorescence.

Finally it must be emphasized that photochemical reactions and chemiluminescence reactions do exist, and that in some reactions at least, radiation must be a factor. At present the evidence seems to indicate that radiation is not important in ordinary thermal reactions.

The radiation hypothesis, like all hypotheses, has helped in the development of science by inspiring new laboratory experiments and further generalizations, and it has exerted a considerable influence on chemical research of the past decade. It has advanced particularly the fields of chemical kinetics and infra-red radiation.

Although three of its four postulates appear to be discredited and there is little support for the fourth, the field is still an attractive one, both for theoretical speculation and experimental exploration.

REFERENCES

- (1) TRAUTZ, Z. anorg. Chem. **102**, 81 (1918) and earlier papers.
- (2) PERRIN, Ann. Phys. [9] **11**, 5 (1919).
- (3) W. C. McCLEWIS, J. Chem. Soc. **113**, 471 (1918) and earlier papers.
- (4) Symposium on Newer Theories of Catalysis, Trans. Faraday Soc. **17**, 500 (1922).
- (5) Symposium on Photo-chemistry. Trans. Faraday Soc. **21**, 438-656 (1925).
RICE 494, ORNSTEIN 504, HIRST AND RIDEAL 508, ROY 512. Discussion 515.
- (6) HARNED, J. Franklin Inst., **196**, 181 (1923).
TAYLOR, H. S., Treatise on Physical Chemistry, D. Van Nostrand Co., New York (1924), p. 1030-52 and 1270-75.
KISS, Chem. Weekblad **20**, 585 (1923).
- (7) PERRIN, "Les Atomes" (1913).
- (8) RICE, J. Am. Chem. Soc. **46**, 2405 (1924).
- (9) DANIELS AND JOHNSTON, J. Am. Chem. Soc. **43**, 53 (1921).
- (10) WHITE AND TOLMAN, J. Am. Chem. Soc. **47**, 1240 (1925).
- (11) HUNT AND DANIELS, J. Am. Chem. Soc. **47**, 1602 (1925).
- (12) HIRST, J. Chem. Soc. **127**, 657 (1925).
- (13) RICE AND GETZ, J. Phys. Chem. **31**, 1572 (1927).
- (14) D. F. SMITH, J. Am. Chem. Soc., **47**, 1862 (1925).
- (15) HINSHELWOOD AND HUTCHINSON, Proc. Roy. Soc. London **111** A, 245 (1926).
- (16) HINSHELWOOD AND THOMPSON, Proc. Roy. Soc. London **113** A, 221 (1926).
- (17) HINSHELWOOD, Proc. Roy. Soc. London **114**, 84 (1927).
- (18) HINSHELWOOD AND ASKEY, Proc. Roy. Soc. London **115**, 215 (1927).
- (19) RAMPSBERGER, (a) J. Am. Chem. Soc. **49**, 912, (1927); (b) J. Am. Chem. Soc. **49**, 1495 (1927).
- (20) D. F. SMITH, J. Am. Chem. Soc. **49**, 43 (1927).
- (21) LINDEMANN, Trans. Faraday Soc. **17**, 598 (1922).
- (22) RODEBUSH, J. Am. Chem. Soc. **45**, 606 (1923).
- (23) HIRST AND RIDEAL, Proc. Roy. Soc., London, **109** A, 526, (1925).
- (24) HIBBEN, Proc. Nat. Acad. Sci. **13**, 626 (1927).
- (25) SMITH AND LOOMIS, Meeting Am. Chem. Soc., Detroit, Mich. Sept. (1927).
- (26) J. J. THOMSON, Phil. Mag. **47**, 337 (1924).
- (27) CHRISTIANSEN AND KRAMERS, Z. physik. Chem. **104**, 451 (1923).
- (28) LEWIS AND SMITH, J. Am. Chem. Soc. **47**, 1508 (1925).

- (29) TOLMAN, J Am. Chem Soc **47**, 1524, (1925).
- (30) HINSHELWOOD, Proc. Roy Soc , London, **113 A**, 230 (1926).
- (31) ROY, Proc Roy Soc., London, **110 A**, 543 (1926)
- (32) CHRISTIANSEN, Proc. Cambridge Phil. Soc., **23**, 438 (1926).
- (33) J. J. THOMSON, Phil. Mag [7] **3**, 241 (1927).
- (34) FOWLER AND RIDEAL, Proc. Roy. Soc London **113 A**, 570 (1927).
- (35) HINSHELWOOD, Proc. Roy. Soc , London **114 A**, 84 (1927).
- (36) HINSHELWOOD AND ASKEY, Proc. Roy. Soc , London, **116 A**, 162 (1927).
- (37) RICE AND RAMPSBERGER, J Am. Chem. Soc , **49**, 1617 (1927).
- (38) LUECK, J. Am. Chem Soc **44**, 757 (1922)
- (39) NORRISH, Nature **119**, 123 (1927); J. Chem Soc. 761 (1927)
- (40) BUSSE AND DANIELS, J. Am Chem Soc. **49**, 1257 (1927)
- (41) ARRHENIUS, Z physik Chem **4**, 226 (1889).
- (42) WULF, Proc Nat. Acad Sci **12**, 129 (1926).
- (43) DAVID, Proc. Roy. Soc London **108 A**, 617 (1925)
- (44) TOLMAN, J Am Chem Soc **42**, 2506 (1920)
- (45) MARCELIN, Ann Phys [9] **3**, 129 (1915).
- (46) RICE, Brit. Assoc. Rep p 397 (1915)
- (47) HINSHELWOOD, Chem Reviews **3**, 247 (1926).
- (48) DUSHMAN, J. Am Chem. Soc. **43**, 397 (1921).
- (49) LANGMUIR, J Am. Chem. Soc. **42**, 2190 (1920)
- (50) TOLMAN, J Am. Chem Soc **47**, 1549 (1925).
- (51) LEWIS AND McKEOWN, J Am. Chem Soc **43**, 1288 (1921)
- (52) DANIELS AND JOHNSTON, J Am Chem. Soc **43**, 73 (1921)
- (53) DANIELS, J Am Chem Soc. **47**, 2856 (1925)
- (54) GARNER, Phil Mag **49**, 463 (1925).
- (55) WARBURG AND LEITHAUSER, Ann. Physik. **28**, 313 (1909).
- (56) W C McC LEWIS, Trans Faraday Soc. **17**, 585 (1922)
- (57) H. A TAYLOR, J Am Chem. Soc. **48**, 577 (1926)
- (58) DANIELS, J Am Chem. Soc **48**, 607 (1926)
- (59) TOLMAN, J Am. Chem. Soc. **47**, 1549 (1925) Statistical Mechanics with Application to Physics and Chemistry. Monograph Am. Chem. Soc , Chemical Catalog Co , New York, p 278 (1927).
- (60) POLANYI, Z Physik **3**, 31 (1920)
- (61) PERRIN, Compt. rend. **178**, 1401 (1924) and preceding communications.
- (62) TAYLOR AND LEWIS, J. Am. Chem Soc **46**, 1606 (1924)
- (63) CAPPER AND MARSH, J. Am. Chem Soc. **47**, 2847 (1925).
- (64) KUGELMASS AND McQUARRIE, Science **50**, 272 (1924).
- (65) KUGELMASS AND McQUARRIE, Science **52**, 87 (1925)
- (66) DANIELS AND FOSBINDER, Science **52**, 266 (1925).
- (67) DAVID, Phil. Mag. [7] **1**, 334, (1926).
- (68) HASLAM, LOVELL AND HUNEMAN, J Ind & Eng. Chem **17**, 272 (1925).
- (69) BONE AND WESTON, Proc Roy. Soc , London **110 A**, 615 (1926)
- (70) GARNER AND JOHNSON, Phil. Mag. [7] **3**, 97 (1927).
- (71) DAVID, Trans. Faraday Soc **22**, 273 (1926).
- (72) DANIELS, Experimental tests of the Radiation Hypothesis. Symposium on Photo Chemistry, Washington, D. C (1924)
- (73) LEWIS AND MAYER, Proc Nat Acad Sci **13**, 623 (1927)

CATALYTIC ACTIONS*

ERIC K. RIDEAL

Cambridge University, England

The voluminous literature and the frequent discussions of catalysis have made us aware of both the scientific interest and the technical importance of catalytic processes. The subject of catalysis has so grown that it is no longer self-contained. Even the chemical reactions and processes in living tissues may in one sense be termed catalytic. It would be impossible in the course of a brief discussion even to sketch the various ramifications into which catalytic processes intrude and it therefore seems most suitable to discuss in some detail the various views which have been expressed with respect to the mechanism of the simpler reactions. Even in this limited section of the subject there is ample provocation to discussion, and I assume that it is the function of the opening address at an extended meeting of this kind to provoke and initiate discussion.

CONTACT CATALYSIS

If we direct our attention to even one simple case we will find that we are only at the beginning of a proper understanding of molecular reactions. On examination of the processes of hydrogenation and dehydrogenation of such substances as ethylene or aldehyde and their products we note that over a great range of pressures the process of dehydrogenation is a zero order reaction whilst in hydrogenation there is an optimum ratio to the partial pressures of the reactants. For a number of dehydrogenating reactions the critical energy increment is constant provided that the surface remains unchanged.

* An address at the opening of the session on catalysis at the first Institute of Chemistry of the American Chemical Society, State College, Pennsylvania, July 14, 1927.

I can refer to the value of $\epsilon = 22,000$ cal./gram mol for the dehydrogenation of a series of alcohols for copper, to the values of $\epsilon = 9800$ for nickel, $\epsilon = 15,700$ for palladium, $\epsilon = 19,000$ for platinised charcoal and $\epsilon = 18,500$ for platinum on charcoal in a variety of dehydrogenating processes such as the decomposition of hexahydrobenzene, decahydronaphthalene and piperidine. If we include the decomposition of formic acid we find $\epsilon = 22,000$ for platinum black, $\epsilon = 31,000$ for silver, $\epsilon = 25,000$ for gold and the exceptional value of $\epsilon = 39,000$ for palladium.

Whilst a large variety of metals are available for dehydrogenating processes it has been suggested that limits are imposed only by the size of the atomic radius which must lie between 1.23 and 1.39 Å.¹ We may note in addition that the hydrides are unstable and are more akin to LiH rather than to HCl. In addition we know that in general the chemical changes produced at the catalyst surface, at not too elevated temperatures, are confined to certain localized areas or patches, but at high temperatures all or a large fraction of the total surface is active. These low-temperature active patches are associated with instability of surface, large energy changes, and long lives of adsorbates. These are the experimental facts and any theoretical treatment of the mechanism must embrace these.

It will be, I think, not disadvantageous to give a brief outline of the views that have been put forward to explain the reactions and to enquire how far these alternative suggestions can be reconciled with all the experimental data that have been established. Faraday's original view involved two assumptions, a multimolecular adsorbed film and a diminution in forces normally preventing chemical action. The existence of multimolecular films, except in gel-like materials (where film-formation can scarcely be said to occur) is now known to be a very rare occurrence and we are confident that reaction, except perhaps in combustive and chain processes takes place at the

¹ It is interesting to note the following sequence in the data of Remy and Schaefer (*Z. anorg. allgem. Chem.* **136** (1924) for hydrogen-oxygen combination on (a) H₂-treated metals: Os > Pt > Pd > Rh (increasing σ); (b) O₂-treated metals: Ir > Pd > Pt (decreasing σ)

surface of the catalyst. A union of the adsorbate with adsorbent would, by rendering the molecules more complex, effect automatically an increase in the fraction of the collisions causing activation as has recently been shown by Christiansen, Hinshelwood, Fowler and myself. This union, to follow the early conceptions of De le Rive and the more recent views of Sabatier, is identical with that in forming definite chemical compounds. A modification of this view has been expressed in that a number of different types of compounds or hydrides are formed, each one capable of effecting different hydrogenations such as of benzene, a double bond or nitrobenzene. Some of the formulas proposed based upon bulk analyses are of course somewhat extraordinary. I do not think this concept of activation by collisions of the second kind of complex molecules formed by union of adsorbate and adsorbent effected by either adsorbate or substrate molecules can be regarded as adequate; for, *per se* it gives no explanation for the existence of active and inactive areas nor for the variation in the critical energy increments for different substrates.

The simplest hypothesis to adopt, if one does not seek for a formulation of the type of union between adsorbent and adsorbate is to assume adsorption of the molecule and that on suitable portions of the catalyst or at high temperatures the molecule is stretched and as a result of such stretching it undergoes activation. We shall return to this view later. At the limit of such extension dissociation into radicles or atoms is effected. Some investigators take the extreme view that activation only results when this limit is attained. There appears to be but slender evidence for this hypothesis. The much sought after atomic hydrogen in hydrogen palladium has not been discovered, whilst in view of the readiness with which cold metals adsorb hydrogen atoms and their ease of formation at hot surfaces the identification of hydrogen atoms on the plate of a thermionic valve cannot, I think, be regarded as conclusive proof that dissociation has been effected by the cold metal plate. Whilst hydrogen atoms can be identified in a number of surface actions it is difficult to bring evidence to prove that they are

not the residuals of a prior chemical action taking place between excited molecules.

Adsorption by electric forces

A number of valuable suggestions have been made to the effect that adsorptive compounds are formed which possess definite electrical structure. It is not difficult to calculate the heat of adsorption for a gas when it consists of dipolar or linear quadripolar molecules, on the assumption that adsorption is effected by attraction into a surface by the mirror image of the molecular electric moment in the substrate material of definite dielectric constant. We may note, however, that this view does not give good values for the relative heats of adsorption on materials of different dielectric constants² and that the absolute values are small and constant, thus giving no interpretation for the observed variation in the heat of adsorption with the configuration of the surface.

Again we find expressed in the literature that adsorption is a surface phenomenon caused by the existence of positive and

² For example the heats of adsorption of argon, nitrogen, and methane on mica and uniform carbon in ergs per molecule are

	Mica	Carbon
	ergs. 10 ¹¹	
A	2 48	0 48
N ₂	2 23	1 03
CH ₄	3 36	1 11

If the heats of adsorption were due entirely to the loss of potential energy due to the movement of a dipole into an equilibrium position in respect to its mirror image, the ratios of the heats of adsorption on carbon and mica should stand in the ratio

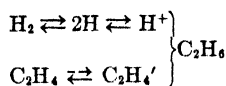
$$\frac{Q_{\text{mica}}}{Q_{\text{carbon}}} = \frac{\frac{K_m - 1}{K_m + 1}}{\frac{K_c - 1}{K_c + 1}}$$

If K_c is taken as large and K_m given the usual value of 6 we obtain for the ratio $\frac{Q_{\text{mica}}}{Q_{\text{carbon}}} = \frac{5}{7}$. It is clear that K_m should have values lying between 3 and 5 to effect agreement between the experimental values and those calculated on this hypothesis.

negative ions already present in the surface of the adsorbent. In its simplest form it is supposed that one atom in the surface of a metal has lost an electron and on adsorption of a hydrogen molecule a molecular hydrogen ion is formed. This idea may be expressed in the form.



Several arguments may be advanced in favor of the view that ions are formed on adsorption on the active portions or on non-active areas at high temperatures. We may note the enormous chemical reactivity of charged ions in the gas phase. Thus N^+ and N^+_2 react readily with hydrogen to form ammonia. CO^+ decomposes to form CO_2 and carbon. Many such reactions have been investigated by means of electron and as Lind has shown by α particle collision. Again we may observe in many surface reactions, especially oxidations on hot wires, the copious emission of positive and negative ions which are imagined to be those which have escaped from the adsorptive layer without combination, on the assumption that the ionic form is the normal state in the layer. Finally it has been stated that with palladium electrodes an ethylene hydrogen cell can be constructed and that the free energy of hydrogenation can be measured with tolerable accuracy by this method, the reactions assumed are



The close parallelism between thermionic emission and chemical reaction at surfaces first noted by Thompson and later by Langmuir has frequently been commented upon. There are likewise experiments indicative of the non-uniform character of the thermionic emission from surfaces as well as a lower work function associated with amorphous surfaces or a layer of amorphous material on a crystalline surface. The Schottky electron emission likewise appears to be confined to spots on the surface; these spots of low work function may be the active patches of low temperature catalysis. In spite of these arguments I do not think that the ionic view is tenable in this simple form. This

can be shown by studying the reactions of hot wires with oxygen or the halogens to form volatile compounds. Langmuir showed that the oxidation rate of tungsten followed a simple unimolecular law. He assumed the existence of a number of free electrons in the metal, some of which would combine with the oxygen molecules on the surface. If we modify this view so as not to accept the view of the existence of numerous free electrons in a metal we find that the critical energy increment is closely approximated to by the difference between the thermionic work function of the metal and the electron affinity of the reacting gas.² It is

² Langmuir showed that the critical energy increment for the surface oxidation of tungsten was 0.6 volts; in our laboratory the value for the analogous process on platinum has been found to be 2.3 volts.

The thermionic work functions for the two metals are

	Platinum		Tungsten
Richardson	5.65v.		
Deininger	5.02	Langmuir	4.25
Wilson	6.0	Smith	4.46
Mean best value	5.5v.	Best value...	4.25

If the critical energy increment E of the surface oxidation be the difference between the work function ϕ of the metal and the electron affinity of molecular oxygen E_r we obtain

$$E = \phi - E_r$$

For Platinum $2.3 = 5.5 - E_r$, where $E_r = 3.2$ volts,

For Tungsten $0.6 = 4.25 - E_r$, where $E_r = 3.65$ volts,

For oxygen there is ample evidence for the existence of O_2' and it is probable, according to Mackay, that the normal molecule has an electron affinity of about 3.5 volts, a value in good agreement with the above.

For halogens it appears from Langmuir's work that they are first thermally dissociated and the atoms then attack the metal. The total energies required for the reaction are for dissociation: Cl_2 , 5.2v at 1000°C.; Br_2 , 2.02; I_2 , 1.51. The electron affinities are: 4.1, 3.8, 3.5.

This gives E values on platinum for the reaction, $X_2 \rightarrow X \rightarrow PtX$, of Cl_2 , 6.6; Br_2 , 3.72; I_2 , 3.51; and on tungsten of Cl_2 , 5.35; Br_2 , 2.47; I_2 , 2.26 volts. The E values for halogen atoms are much smaller, viz.:

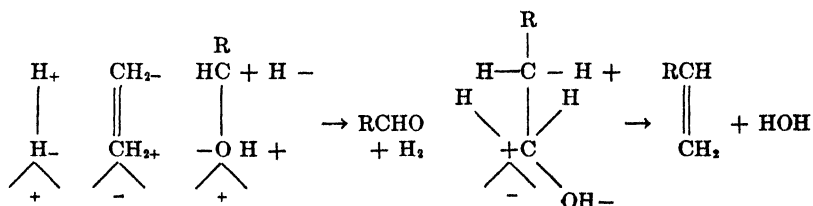
	Cl	Br	I
Pt	1.4	1.7	2.0v
WO	0.15	0.435	0.7v

showing that reaction will take place rapidly even at low temperatures.

It is interesting to note in this connection that Davy in 1807 (Phil. Trans.) showed that negatively charged zinc could not be oxidized and that silver becomes oxidizable on acquiring a positive charge.

clear that all the reacting gas is not present on the surface as ions, but the molecules are only converted to ions after the supply of a certain amount of energy as measured by the critical energy increment. Again, although Thompson many years ago showed that reactions between hydrogen and oxygen on a platinum wire became vigorous at the temperature at which the thermionic emission *in vacuo* became marked, rapid combination can occur at the surface of many metals at very low temperatures. Finally the rate of hydrogen-oxygen combination greatly exceeds the rate of metal oxide formation, an indication that the oxygen need not pass through the oxide form to ensure reaction with hydrogen.

There remains finally one other view which merits some attention, which also involves the existence of or induction of positive and negative charges on the surface of the adsorbate, viz., the polarization of molecules on adsorption and the entry of adsorbate molecules within the double layer on the surface of the adsorbent. These polarized molecules are assumed to be chemically reactive. It is a simple matter to construct the usual kind of picture to indicate the ideas underlying this view, such as:



It seems probable that the readiness with which a gas is adsorbed is dependent on the magnitude of its electric moment, but the mechanism of attraction on metals cannot involve the simple mirror concept. It would appear that in catalytic surfaces the adsorbate must enter the field exerted by the surface⁴ and we

⁴ If the free surface energy of a material σ is due to the electrostatic energy of its surface energy of its surface dipoles it has shown that $\sigma = \frac{9V^2}{80\pi r}$ where V is the potential for photo electric emission. For mercury $V = 4.2$ volts, $r = 1.4 \text{ \AA}$, whence $\sigma = 472$ dynes per cm. (observed $\sigma = 465$).

must have the presence of natural or induced dipoles which are characteristic of the metal or other catalyst and which depend on the texture as well as the nature of the material. Some evidence for this view can be obtained from other sources such as in the behavior of metallic beams impinging on cold surfaces originally examined by Wood, Langmuir and Frenkel. The evidence is fairly conclusive that duplets or small aggregates of atoms on a plane surface not only evaporate less freely than single atoms but also exert a marked local field. It may be observed that if Frenkel's concept of a solid metal as a system in which the valency electrons whilst never free are in continuous movement from atom to atom be correct, we obtain a simple mechanism for the formation of dipoles possessing both electrostatic and electromagnetic moments in pairs of atoms resting on the surface.

Thus the surface where adsorption takes place on a catalyst on this view would consist of fluctuating systems of electric moments of various magnitudes which are dependent on the closeness of atom packing and elevation above the mean surface of the catalyst, and the presence of true promoting atoms.

Molecular deformability

On adsorption a distortion of the adsorbed molecule takes place. This may be an actual expansion, but deformation resulting in an increase in the electric moment P may or may not result in an increase in the value of the mechanical moment J , as it does not necessarily follow that on excitation of dipolar molecules with small P values any corresponding movement in the spacing of the atomic centers takes place. From the following data it will be observed that in some cases J decreases as P increases.

	Moments of		
	N_2 10 ⁻⁴⁰	CO 10 ⁻⁴⁰	S_2 10 ⁻⁴⁰
Unexcited (fine structure of band) . .	3 59	17 83	12.6
Excited	3 33	14.23	13.8
from vapor pressure	2 82 S	3 06 S'	

S is the symmetry number in the expression for the chemical constant

$$i = \lg \cdot \frac{(2\pi m)^{\frac{3}{2}} 8 \pi^2 J R^2}{5 h^3}$$

Again, in the case of sodium chloride Reis has noted on the assumption that the condition of equilibrium of the salt molecule both in the gas phase and in the lattice is given by the expression $\frac{d}{dr} \left(\frac{e^2}{r} - \frac{B}{r^9} \right) = 0$ that r is smaller in gaseous NaCl than in the lattice where the deformation is greater.

The conditions of activation on adsorption thus necessitates a dipole of sufficient electrostatic or electromagnetic moment on the surface, the existence of an electrostatic moment in the absorbate,⁵ and a readiness for the molecule to undergo deformation. We may note that the deformability of a molecule has been defined by Born in the expression $\alpha = \frac{P}{\epsilon}$ where P is the electric moment of the dipole in an electric field of intensity ϵ . The value of α for a reactant should be the factor on which ease of activation is dependent. The classical method for the calculation of α is from the refractivity $\alpha = \frac{3}{4} \frac{n^2 - 1}{n^2 + 2} \frac{M}{\pi N}$ and the value obtained in this manner can in many cases be compared with those determined from the conditions of crystal

⁵ There is a very considerable difficulty in the case of hydrogen. Generally this gas is treated as a simple dipole and numerous workers such as Reiche, Planck, Ehrenfest, Schrödinger, and Kemble have derived specific heat curves on the basis of various *a priori* probabilities and the exclusion of certain quantum states, and from these curves have obtained an estimate of J , the rotational heat, $\frac{C_r}{R}$, being a function of the temperature. The best value with the observed data is $J = 2.095 \cdot 10^{-41}$ gr. cm² (Reiche); but this does not agree with data, e.g., from pressure, $J = 0.1 \cdot 10^{-40} S$, where S is the symmetry number, — or with the value from the line spectrum, $J = 4.5 \cdot 10^{-41}$. Arbitrary suppression of specially chosen states (Schrödinger) improves matters but the process appears artificial. Dennison's more recent hypothesis of hydrogen existing as a mixture of two gases with different properties although not very intelligible as a physical conception, does in fact give an excellent value for J .

equilibrium, from electrostriction or from the temperature coefficient of the dielectric constant.

We find for example the following values for α in the two oxides of carbon:

	α	P
CO	{from structure 1 10 10 ⁻²⁴	1.18 10 ⁻²⁰
	{from refractivity 1 73 10 ⁻²⁴	
CO ₂	{from structure 0.46 10 ⁻²⁴	1.42 10 ⁻¹⁹
	{from refractivity 1 43 10 ⁻²⁴	

It is clear that carbon dioxide possesses a large electric moment and is thus readily adsorbed but its deformability, and thus its ease of excitation is on the whole somewhat less than carbon monoxide.

The various mathematical problems connected with the question of dielectric constant and temperature such as were originally formulated in the Langevin-Debye expression for susceptibility and temperature seem to be most readily developed by methods of the De Broglie-Schrodinger wave mechanics.

Ultimate disruption of a dipole naturally leads to the formation of ions, not of atoms, and thus the ions observed in surface reactions, especially at high temperatures, may be not only, as Haber suggests, the by-products of a chemical reaction, nor, as Brewer believes them to be, a few which have escaped from the cycle by which the whole surface action operates, but may actually consist of a few molecules which have been ejected by thermal agitation after adsorption on suitable portions of the surface where the gain in potential energy on adsorption is higher than the limiting lower value at which the surface action may progress. It is indeed possible that the character of the surface action changes when the temperature is sufficiently elevated for this ejection of ions to proceed, and we obtain a transition from a surface to a chain reaction, a phenomenon readily noted in the hydrogenation of oxygen or even better of bromine at a platinum surface.

One other point I might allude to in this consideration of electric moments. By examining the modification of an air liquid

interface potential by the introduction of a unimolecular film on the surface Guyot and later Frumkin following Guyot's method, as well as Smythe, from determination of the refractive index, came to the conclusion that electric moments could be assigned to a number of so-called polar groups $-\text{OH}$, $-\text{COOH}$ and the like. Now the very interesting property of residual charge, as well as the energy loss in dielectrics under the influence of an alternating field appears to be associated with substances possessing electric moments or capable of acquiring such when minute displacements take place. This energy is not released immediately but only slowly.

The observations suggest that the act of adsorption proceeds isothermally and the adsorbed molecule is still in possession of the potential energy lost in falling into the surface. Such adsorbed molecules are thus potentially reactive until they loose their potential energy with the evolution of heat, the heat of adsorption. That this thermal evolution occurs but slowly when the surface is sparsely covered is an indication that the means of conversion of the potential energy into kinetic is either by collision with adsorbed neighbors as imagined by Garner or by isochronic induction in its neighbors as indicated from Perrin's work on fluorescence. It may be noted, however, that on adsorption of gases on substrates such as carbon, the substrate undergoes expansion and the energy changes of such should be investigated before definite conclusions can be drawn from heat evolution measurements.

REACTIONS IN SOLUTIONS

One of the most interesting developments in reactions in homogeneous systems has been the regularity with which reactions appear to conform to the ideas of Bjerrum and Brønsted on the mechanism of such actions. You will remember that with the breakdown of the simple formulation of the reaction velocity in solution as a function of the concentrations of the reactants, i.e.

$$\frac{dx}{dt} = K (A) (B)$$

the activity concept of G. N. Lewis was introduced by McC. Lewis and Scatchard in the form

$$\frac{dx}{dt} = K f_A(A) f_B(B)$$

where f_A f_B are the activity coefficients of the reactants. Close investigation showed that whilst this formulation appeared to be correct for certain reactions, e.g., the decomposition of hydrogen peroxides by hydrochloric acid it was certainly not generally true. Brönsted and Bjerrum suggested that the mechanism of reaction involved the formation of an "association complex" or "quasi-compound." This complex it was suggested could either revert back to the original reactants with which it was in mass equilibrium or could undergo reaction at a rate proportional to its concentration. We can thus formulate our reaction mechanism in the form

$$\frac{dx}{dt} = K(AB)$$

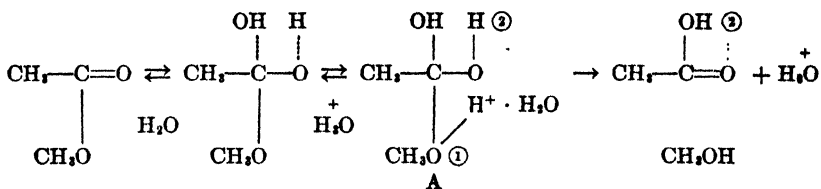
Since by the mass law $\frac{f_A(A) f_B(B)}{f_{AB}(AB)} = K$ the equilibrium constant, we obtain

$$\frac{dx}{dt} = K \frac{f_A(A) f_B(B)}{K f_{AB}}$$

an expression different from that suggested by W. C. McC. Lewis and Scatchard, but closely approximating to it if f_{AB} the activity coefficient of the complex is the activity coefficient of an uncharged complex for as Brönsted has shown experimentally and as the work of Milner, Debye and Hückel has indicated the magnitudes of the activity coefficients are dependent chiefly on the charge on the particular molecular species.

The recent investigations of Dawson can be shown to conform to this concept, but in addition we can draw the inference from his work that in reactions catalysed by oxonium and hydroxyl ions specific catalytic activities can be ascribed to all oxonium ion "donators" and oxonium ion "acceptors" to adopt

Wieland's nomenclature. Thus in the case of catalysis by means of dilute aqueous acetic acid specific catalytic activities may be ascribed to the following, the oxonium ion donors H_2O , CH_3COOH , H_3O^+ and the oxonium ion acceptors OH' , $\text{CH}_3\text{COO}'$. The original dual theory of Snethlage has by now undergone a very considerable extension. It is not, I think, necessary to postulate in say the hydrolysis of an ester by a dilute acid that an oxonium ion is "donated" at the identical moment that another one is "accepted" but the reactions present in interesting features as can be noted from the following sequence.



It will be noted that the reaction involves the addition of one hydrogen ion (1) and the removal of another (2) effected by the passage of an electron from (2) to (1). We must remember that the complex (A) is the critical complex or quasi compound and is consequently capable of reversion to ester and oxonium ion. Whilst we may regard the action of hydrolysis as a species of conduction or even of electrolysis if this is not straining the views of Professor Armstrong unduly, we do not yet know what conditions the forward or reverse change, if hydrolysis occurs at some crisis in the molecular movements or if the supply of the energy of activation affects not the concentration of quasi-compound but its fragility. It is clear, however, that if either of these views be correct the quasi compound (a) must have an objective existence even if present in subanalytical amounts and the two steps in the process, the addition of an oxonium ion at (1) and the loss of one at (2) cannot be simultaneous processes.

Molecular induction

Perrin in the early development of the radiation theory of chemical action suggested that excited molecules would loose their energy of excitation not only by super elastic collisions as

proposed by Klein and Rossland but also by isochronic induction. More recently attempts have been made in various directions to calculate the energy available when a molecule enters the field either of another molecule of definite electric moment, or the field of an ion. Such investigations have been carried out notably by Sir J. J. Thomson, Merritt and W. Taylor, with the view of examining the importance of ions and polar molecules on the rates of chemical action. The former has shown that the rate of collision between reactants A and B can be increased very considerably when either reactant is replaced by an ion, e.g., A^+ or by a compound such as bA where bA represents a primary complex between a polar molecule b and the reactant A . All such collisions are, as we know, not effective in promoting chemical action and it is important to enquire whether the critical energy increment of these reactions may be modified as well as the collision frequency by the alteration in field produced by a molecule of definite electric moment or by one ion. It is assumed that when a dipolar molecule B enters the electric field of a polar molecule b or of an ion A^+ it acquires potential energy due to the orientation of its axis to the lines of force, the moment of inertia J as well as the electric moment of B is altered and a quantity of energy ϵ will either be liberated or absorbed in this process. Merritt assumes that this energy contributes to the energy of activation. The following two simple calculations show the relative magnitudes of the energy changes effected in the two cases.

We can, on the assumption of a molecule entering the field at a distance some four molecular radii away, equate the field strength to $\frac{\mu}{(4\sigma)^2}$ or if the molecule be water and $\mu = 1.10 \cdot 10^{-18}$ E.S.U. $\sigma = 2.3 \text{ \AA}$, the field strength will be $1.3 \cdot 10^3$ E.S.U., or $3.85 \cdot 10^5$ volts per centimeter. The maximum potential energy of a molecule of moment μ_1 in this field will be $\mu_1 \times 1.3 \cdot 10^3$ E.S.U. For a molecule like hydrogen chloride where $\mu_1 = 3.10 \cdot 10^{-18}$ we obtain $= 3.84 \cdot 10^{-16}$ ergs or only some 55 calories per gram molecule. This small value clearly negatives Cathala's view on the catalytic influence of water in the hydrogen chloride combination, but at the same time it shows that intense fields may be produced if

the water molecule is more clearly associated with the reactant than at a distance of 4σ as assumed in this calculation. At a distance of one radius ϵ has already risen to 3000 calories per gram molecule.

If the polar water molecule be replaced by an ion we shall obtain round the ion a cluster of molecules as visualised by Langevin, Sir J. J. Thomson, Erikson and Loeb and the average potential energy of a molecule in the first shell at a distance 2.3 \AA will be some 12,000 calories per gram molecule.

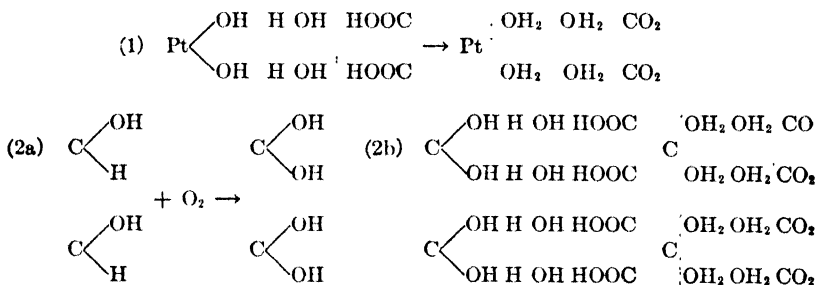
It will be observed that the possible gain in potential energy of a molecule entering the field of an ion is remarkably large and consequently marked catalytic effects might be anticipated for reactions in which ions are involved.

Oxidative processes

Attention is again being devoted to the mechanism of slow and rapid combustion and in view of the technical importance of explosive reactions on the one hand and of autoxidation and slow combustion on the other this renewed vigour exhibited in this field during recent years is very welcome.

It appears certain that a number of oxidations in solution are brought about on surfaces, we may mention the oxidation of oxalic and other acids on charcoal, and of benzaldehyde and turpentine on glass, pumice and similar surfaces. By the method of selective poisoning it can be shown that but a small portion of the total surface of say powdered glass is catalytically active. Many substances exert what Moureaux terms an antioxygenic action by selective adsorption on these surfaces; this is true for iodine, diphenylamine and even organic acids. These oxidation reactions involve in many cases the formation of a definite and relatively stable peroxide which in the case of aldehyde peroxide undergoes reaction in the homogeneous liquid phase to the acid. It is, I think, almost certain that prior to the formation of the stable peroxide there exists another and more reactive form of peroxide which is capable of propagating chains in the sense suggested by Christiansen and Kramers. On surfaces at any rate it is very probable that typical inter-face reactions take place the combustion of one molecule of

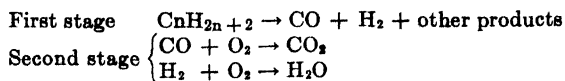
aldehyde on a glass surface results in the "clean up" of many molecules on the surface, by a spreading of the combusive process. On charcoals at least five different types of surface can be identified. There exists a small area which is autoxidizable and recently Ward has shown that these autoxidizable areas are identified as the areas in which Blench and Garner noted large values for the heats of adsorption of oxygen. In addition there exists catalytically active areas on which various oxidative processes as shown by Warburg and Miss Wright can take place. It is an interesting fact that the rates of oxidation on this surface for various substances at their respective optimum concentrations run parallel to the ease with which they reduce the oxygen overpotential at a platinum anode. Finally there exist two promoted areas, one associated with a complex containing both carbon and iron and another extremely active area associated with a complex containing iron, carbon, and nitrogen. These curious reactions call attention to a possible connection with the suggestions of Professor Armstrong concerning the nature of reactions as a species of electrolysis, these may be depicted formally by the following equations:



It must not be forgotten, however, that the existence of a carbon "perhydrol" similar to that postulated for platinum rests on somewhat slender experimental basis.

In explosive oxidations affairs are certainly more complex. We may observe in rapid oxidations such as of phosphorus vapor or of benzaldehyde vapor at high temperatures, the emission of radiation. These reactions are undoubtedly chain reactions and since in several cases emission of radiation can actually be

observed it is probable that "quantum" chain reactions do occur. Further work is still required I think to demonstrate the existence of the chains formed by collision of the second kind between excited products and reactants as postulated by Christiansen and Kramers. Whilst it has been suggested especially by Egerton that peroxides play an important part in explosive reactions the view that these are formed after primary dissociation of a saturated hydrocarbon into an unsaturated one and hydrogen has taken place does not rest upon such a firm foundation and it is equally probable that peroxide formation can occur on reaction between excited molecules of the saturated hydrocarbon and oxygen. Egerton considers that the conditions determining peroxidation of oxidizable substances likewise determine the conditions for detonation. This view is at variance with that suggested from examination of the spectrum emitted during explosions of gaseous mixtures. Whilst it has long been suggested that a zone of electrons or radiation travels in front of the explosive wave and acts as an exciting agent for the inherent gas yet Garner concludes that the ionization produced during gaseous reactions is no greater than that which would normally occur due to thermal ionization at elevated temperature on the explosion. Affairs appear to be otherwise with respect to the quantity and nature of the infra red radiation emitted. In dry gas mixtures, e.g., carbon monoxide the emission of such radiation is far greater than in wet mixtures or in mixtures which form water on combustion. The fine structure of the band spectrum indicates that excited water molecules containing both rotational and vibrational energy are the sources of the radiation. Garner regards water as an energothermic catalyst in that it absorbs radiation otherwise emitted and facilitates the distribution of energy in the system. We might represent the sequence of events somewhat as follows:



Antiknock compounds are assumed to slow down the rate of water formation and thus permit of an increased dissipation of energy in the form of radiation.

REFERENCES

- DEBYE. *Physik. Z.* **21**, 178 (1920).
 GICKLHORN: *Z. physik. Chem.* **123**, 344 (1926).
 MERRITT: *Phys. Rev.* **28**, 684 (1926).
 MERRITT: *J. Optical Soc. Am.* **12**, 613 (1926).
 MILLIKAN AND EYRING: *Phys. Rev.* **27**, 51 (1926).
 RICHARDSON AND YOUNG: *Proc. Roy. Soc. (London)* **107A**, 377 (1925).
 THOMSON: *Phil. Mag.* **47**, 337 (1924).
- Electric moments*
 COMPTON AND ZAHN: *Phys. Rev.* **23**, 781 (1924).
 ERRARA: *J. phys. radium*, **6**, 390 (1925).
 SÄNGER: *Physik. Z.* **27**, 556 (1926).
 WOLF: *Physik. Z.* **27**, 588 (1926).
- Wave mechanics*
 VAN VLECK: *Phys. Rev.* **29**, 727 (1927).
- Deformability, CO*
 MATOSSI. *Z. Physik* **40**, 1 (1926).
 WEIGT: *Physik. Z.* **22**, 643 (1921).
- Deformability S₂*
 TAYLOR AND RIDEAL: *Proc. Roy. Soc. (London)* **115A**, 589 (1927).
- Deformability, CO₂*
 BARKER: *Astrophys. Z.* **55**, 391 (1922).
 BORN AND HEISENBERG: *Z. Physik.* **23**, 388 (1924).
 EUCKEN: *Z. physik. Chem.* **100**, 159 (1922).
 KLIEFOTH. *Z. physik.* **39**, 402 (1926).
 REIS: *Z. physik* **1**, 294 (1920).
- Energies of activation*
 ZELINSKY AND BALANDIN: *Z. physik. Chem.* **126**, 270 (1927).
- Adsorption of dipoles*
 EUCKEN. *Z. Elektrochem.* **28**, 6 (1922).
 JACQUET: *Fortschritte Chem., Physik, physik. Chem.* **18**, 437 (1925).
- Adsorption as ions*
 BREWER: *Proc. Nat. Acad. Sci.* **12**, 560 (1926).
 ERIKSON: *Phys. Rev.* **26**, 625 (1925).
 SCHMIDT: *Z. physik. Chem.* **118**, 193 (1925).
- Adsorption in pairs*
 CASSELL: *Physik. Z.* **28**, 152 (1927).
 MAGNUS: *Z. anorg. allgem. Chem.* **158**, 67 (1926).
- Adsorption as atoms*
 FRYLING: *J. Phys. Chem.* **30**, 818 (1926).
 GAUGER: *J. Am. Chem. Soc.* **46**, 674 (1924).
 KISTIAKOWSKY: *J. Phys. Chem.* **30**, 1356 (1926).
 WOLFENDEN. *Proc. Roy. Soc. (London)* **110A**, 464 (1926).
- Homogeneous reactions*
 DAWSON AND CO-WORKERS: *J. Chem. Soc.* **129**, 2282, 2872, 3166 (1926); **131**, 213, 458, 756 (1927).

THE STRUCTURE OF ATOMS AS A PERIODIC PROPERTY AND ITS RELATION TO VALENCE AND ION-FORMATION*

ARTHUR A. NOYES AND ARNOLD O. BECKMAN

Gates Chemical Laboratory, California Institute of Technology

INTRODUCTION

The great development of the knowledge of the electronic structure of atoms which has resulted during the last two or three years from spectroscopic investigations has made it possible to draw a much more complete picture of the atoms of the elements of the first two short periods, and in less degree of those of the first long period of the periodic system. Namely, it is now possible in the case of most of these atoms and of many of the ions derived from them to specify not only the number and the quantum-states of the electrons in each of the shells of the atom, but also to evaluate the energies attending the removal of its electrons from the various quantum levels. Tables were published in 1924 by Main-Smith (19) and Stoner (20) showing the now accepted distribution of the electrons among the various quantum levels; and these tables have since been extended (21)¹ and reproduced by others (22). The corresponding values of the energies of the removal of the electrons have, however, not been compiled in connection with such tables. It seemed therefore well worth while to collate these values, which are scattered through the recent literature, and to present in the form of a

* Contribution No. 177 from the Gates Chemical Laboratory. The more general features of this paper were presented in a presidential address before the Pacific Division of the American Association for the Advancement of Science at its Reno meeting on June 22, 1927; and a much briefer form of this article was published in *Proc. Nat. Acad. Sci.*, 13, 737-743 (1927).

¹ Thus by McLennan, McLay, and Smith, *Proc. Roy. Soc. London*, (A) 112, 76 (1926), who added the quantum characteristics of the basic spectral term for each atom, especially its inner quantum number.

chart a graphic picture of the present knowledge of the structure of atoms, showing these energies in connection with the quantum states of the atoms.

Desirable as it would be to do so, it did not seem practicable at the present time to extend the chart and discussion of the results beyond zinc, because the interpretation of the spectroscopic data of the higher elements is still so fragmentary that energy values are seldom available, except for the more loosely bound electrons.

In the assembling and interpretation of the spectroscopic data we have been greatly assisted by the advice of Professor Ira S. Bowen of this Institute, by whose experimental researches, made in association with Dr. R. A. Millikan, the data for the chart have in large measure been made available.

THEORETICAL CONSIDERATIONS

The structure data here presented, since they largely relate to the outer shell of electrons, are based almost wholly on spectroscopic studies. From empirical studies there resulted, even prior to the development of the quantum theory, a remarkable classification into series of the complex spectral lines; but the quantum theory suggested further principles which within the last few years have made almost complete the methods of ordering spectra² though a great deal of work must still be done in applying these methods to the individual elements.

It is to be noted that the data to be presented below of the energies of removal of the electrons from atoms are not based on any hypothetical considerations relating to the dynamics of the atom, such, for example, as the Bohr theory of orbits; and that

² A good elementary presentation of the subject has been given by Goudsmit in 1927 in a dissertation entitled "The Atomic Models and the Structure of Spectra." This is now being translated and revised, with the coöperation of the author, by Dr. Linus Pauling of California Institute of Technology. For readers familiar with earlier quantum treatments of spectra, as given in Sommerfeld's "Atomic Structure and Spectral Lines," the following references may be found helpful: Sommerfeld, "Three Lectures on Atomic Physics," Methuen and Co., Ltd., London, 1926; F. Hund, "Linienpektren und Periodisches System der Elemente," Julius Springer, Berlin, 1927.

they will remain valid even though new conceptions as to atom mechanics be developed. For, so long as the basic quantum principle $E = h\nu$, stating that the energy E of each quantum of a radiation is proportional to its frequency ν , is accepted, the so called "terms," which are the values whose differences give directly the frequencies of observed spectral lines, will retain their significance as energies.

As is well known, the electrons of the various atoms have been arranged in a series of shells, each of which, with increasing atomic number or nuclear charge, becomes filled when a certain maximum number of electrons has been added. Thus, the helium atom has a shell of 2 electrons, known as the 1 or K shell; the neon atom has in addition a second shell of 8 electrons, known as the 2 or L shell; the nickel atom has also a third shell of 18 electrons, known as the 3 or M shell; etc. The numbers 1, 2, 3, 4 . . . correspond to the principal or total quantum number n . It was, moreover, recognized by Bohr (23) in 1921 that all of the electrons within one of these main shells are not equivalent; and he proposed a classification into subshells, which he assumed to contain equal numbers of electrons. Thus the 8 electrons of the L shell were considered to form two subshells each containing 4 electrons, and the 18 electrons of the M shell to form three subshells each containing 6 electrons. Main-Smith (24) in 1924 and Stoner (20) some months later showed that the chemical facts and the spectra, respectively, are much better accounted for by assuming that the shells consist of subshells containing the following numbers of electrons: the K shell, 2 electrons; the L shell, a subshell of 2 and one of 6 electrons; and the M shell, subshells of 2, 6, and 10 electrons.*

The recent developments in the interpretation of the spectra of complex atoms by Russell and Saunders (25), Pauli (26), Heisenberg (27), and Hund (28) have completely confirmed the scheme of Main-Smith and Stoner. They have, moreover, made it possible to derive from the optical spectrum of any atom or ion the energies of removal of the electrons from its outer subshells.

* For a table showing the electron groups for all the atoms, see Andrade (22).

In order to describe the state of the electrons in the subshells, another series of integral quantum numbers is used. These were originally called the azimuthal quantum numbers k ; for this number, according to the Bohr orbital theory, determined, when considered in relation to the total quantum number, the eccentricity of the orbit of the electron. Its value, which ranges from unity up to that of the total quantum number, is indicated by attaching it as a subscript to the total quantum number. Thus 2_2 , 3_3 , 4_4 denote circular orbits; 2_1 , 3_2 , 4_3 the elliptical orbits of least eccentricity; 3_1 , 4_2 elliptic orbits of greater eccentricity. More recent quantitative developments of the quantum theories of spectra and of the new wave-mechanics have led physicists to replace the azimuthal quantum numbers by values that are uniformly smaller by one unit, these values being denoted by the symbol l . Spectroscopists employ a different notation, derived from the designations of spectral series, in which conventional letters are used, instead of integral numbers. The relations between the three systems of notation are as follows:

Values of k	1	2	3	4	5	6
Values of l	0	1	2	3	4	5
Spectroscopic notation	s	p	d	f	g	h

Thus the states denoted by 3_1 , 3_2 , and 3_3 using the azimuthal numbers are represented in spectroscopic notation by $3s$, $3p$, and $3d$ respectively. For the values in table 1, which were compiled mostly from spectroscopic sources, the spectroscopic notation is retained; but in the charts and throughout the discussion the azimuthal notation is employed.

THE QUANTUM STATES AND ENERGIES OF REMOVAL OF THE ELECTRONS

In table 1 are given, for the first 30 elements, the number of electrons in the outer subshells of the neutral atoms and of the various ions resulting from them, together with their energies of removal. Isolated values for a few higher elements are also included.

By the energy of removal of an electron is meant the energy

TABLE I
Quantum states and energies of removal of electrons

ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE*	ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE*
		Number	State	E	\sqrt{E}				Number	State	E	\sqrt{E}	
1	H	1	1 s	1.000	1.000	—	7	N ²⁺	2	2 s	4.16	2.04	Est.
2	He	2	1 s	1.803	1.343	1	8	N ³⁺	2	2 s	5.43	2.33	Est.
	He ¹⁺	1	1 s	4.000	2.000	—		N ⁴⁺	1	2 s	7.193	2.682	4
3	Li	1	2 s	0.397	0.630	2		O	4	2 p	1.002	1.001	2
		2	1 s	4.90	2.21	Est.†			2	2 s	2.094	1.447	3
	Li ¹⁺	2	1 s	5.6	2.37	Est.		O ¹⁺	3	2 p	2.585	1.608	2
	Li ²⁺	1	1 s	9.000	3.000	—			2	2 s	3.35	1.83	Est.
4	Be	2	2 s	0.702	0.838	2		O ²⁺	2	2 p	4.047	2.012	2
		2	1 s	8.5	2.9	Est.			2	2 s	5.15	2.27	Est.
	Be ¹⁺	1	2 s	1.340	1.157	2		O ³⁺	1	2 p	5.687	2.385	2
		2	1 s	9.6	3.1	Est.			2	2 s	6.30	2.51	Est.
	Be ²⁺	2	1 s	11.4	3.38	Est.		O ⁴⁺	2	2 s	8.07	2.84	Est.
	Be ³⁺	1	1 s	16.000	4.000	—		O ⁵⁺	1	2 s	10.144	3.185	4
5	B	1	2 p	0.616	0.785	2	9	F	5	2 p	1.248	1.117	2
		2	2 s	0.886	0.941	3			2	2 s	2.74	1.66	3
	B ¹⁺	2	2 s	1.787	1.337	2		F ¹⁺	4	2 p	2.386	1.545	2
	B ²⁺	1	2 s	2.791	1.670	2			2	2 s	4.00	2.00	Est.
	B ³⁺	2	1 s	19.3	4.39	Est.		F ²⁺	3	2 p	4.62	2.15	Est.
6	C	2	2 p	0.835	0.914	2			2	2 s	6.10	2.47	Est.
		2	2 s	1.18	1.09	3		F ³⁺	2	2 p	6.40	2.53	Est.
	C ¹⁺	1	2 p	1.794	1.339	2			2	2 s	7.34	2.71	Est.
		2	2 s	2.2	1.48	Est.		F ⁴⁺	1	2 p	7.56	2.75	Est.
	C ²⁺	2	2 s	3.360	1.833	2			2	2 s	9.24	3.04	Est.
	C ³⁺	1	2 s	4.744	2.178	2		F ¹⁻	6	2 p	0.52	0.72	Est.
7	N	2	1 s	29.2	5.40	Est.	10	Ne	6	2 p	1.587	1.260	1
		3	2 p	1.070	1.035	2			2	2 s	3.57	1.89	Est.
		2	2 s	1.62	1.27	3		Ne ¹⁺	5	2 p	3.025	1.738	15
	N ¹⁺	2	2 p	2.183	1.478	2			2	2 s	4.84	2.20	Est.
		2	2 s	2.76	1.66	Est.	11	Na	1	3 s	0.378	0.615	1
	N ²⁺	1	2 p	3.486	1.867	2			6	2 p	2.80	1.673	3

* References for these tables have been placed at the end of the article, before the general references.

† Est. = estimated. The values given for negative ions, which were derived by extrapolation by the Millikan-Bowen principle, should be regarded as only rough indications.

TABLE 1—Continued

ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE	ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE
		Number	State	E	\sqrt{E}				Number	State	E	\sqrt{E}	
11	Na	2	2 s	4.71	2.17	Est.	16	S	4	3 p	0.761	0.873	1
	Na ¹⁺	6	2 p	3.47	1.86	3			2	3 s	1.485	1.219	3
		2	2 s	5.71	2.39	Est.		Si ¹⁺	3	3 p	1.74	1.32	Est.
12	Mg	2	3 s	0.562	0.750	1	17	Si ²⁺	2	3 p	2.37	1.54	Est.
		6	2 p	3.76	1.94	Est.		Si ³⁺	1	3 p	3.479	1.865	8
		2	2 s	6.30	2.51	Est.		Si ⁴⁺	2	3 s	4.93	2.22	Est.
		1	3 s	1.106	1.051	1		Si ⁵⁺	1	3 s	6.473	2.544	9
	Mg ¹⁺	6	2 p	5.04	2.25	3		Si ¹⁻	5	3 p	0.22	0.47	Est.
		2	2 s	6.71	2.59	Est.		Si ²⁻	6	3 p	0.00	0.02	Est.
	Mg ²⁺	6	2 p	5.98	2.45	3							
		2	2 s	7.56	2.75	Est.		Cl	5	3 p	0.96	0.98	3
									2	3 s	1.81	1.35	3
13	Al	1	3 p	0.440	0.663	1	18	Cl ¹⁺	4	3 p	1.66	1.29	3
		2	3 s	0.81	0.90	3		Cl ²⁺	3	3 p	2.93	1.71	16
		6	2 p	5.0	2.24	5		Cl ³⁺	2	3 p	3.50	1.87	3
		2	2 s	8.29	2.88	Est.		Cl ⁴⁺	1	3 p	5.00	2.24	3
	Al ¹⁺	2	3 s	1.342	1.159	1		Cl ⁵⁺	2	3 s	6.55	2.56	Est.
	Al ²⁺	1	3 s	2.092	1.446	1		Cl ⁶⁺	1	3 s	8.400	2.898	4
	Al ³⁺	6	2 p	9.00	3.00	3		Cl ¹⁻	6	3 p	0.33	0.57	Est.
								A	6	3 p	1.159	1.076	10
14	Si	2	3 p	0.761	0.872	3	19	A ¹⁺	2	3 s	2.25	1.50	Est.
		2	3 s	1.00	1.00	Est.			5	3 p	2.00	1.41	15
		6	2 p	7.2	2.70	5							
		2	2 s	10.9	3.3	Est.	20	K	1	4 s	0.319	0.565	1, 18
	Si ¹⁺	1	3 p	1.202	1.096	6			6	3 p	1.90	1.38	Est.
		2	3 s	1.49	1.22	Est.			2	3 s	3.17	1.78	Est.
	Si ²⁺	2	3 s	2.338	1.529	6	21	K ¹⁺	6	3 p	2.34	1.53	17
	Si ³⁺	1	3 s	3.320	1.822	6							
	Si ⁴⁺	6	2 p	12.47	3.53	3		Ca	2	4 s	0.450	0.671	1, 18
15	P	3	3 p	0.982	0.991	1	20		6	3 p	2.72	1.65	Est.
		2	3 s	1.52	1.23	Est.			2	3 s	4.25	2.06	Est.
		6	2 p	9.5	3.08	5		Ca ¹⁺	1	4 s	0.873	0.934	1, 18
		2	2 s	13.7	3.7	Est.	21	Ca ²⁺	6	3 p	3.76	1.94	17
	P ¹⁺	2	3 p	1.462	1.209	7							
	P ²⁺	1	3 p	2.216	1.489	8		Sc	2	4 s	0.49	0.702	11, 18
	P ³⁺	2	3 s	3.53	1.88	Est.	21		1	3 d	0.55	0.76	Est.
	P ⁴⁺	1	3 s	4.778	2.186	9			6	3 p	3.4	1.85	Est.
									2	3 s	5.1	2.27	Est.

TABLE 1—*Concluded*

ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE	ATOMIC NUMBER	ATOM OR ION	ELECTRONS		ENERGY OF REMOVAL		REFERENCE
		Number	State	E	\sqrt{E}				Number	State	E	\sqrt{E}	
1	Sc ¹⁺	1	4 s	0.948	0.974	11, 18	26	Fe	2	4 s	0.578	0.759	18
		1	3 d	1.1	1.05	Est.			6	3 d	0.81	0.90	Est.
	Sc ²⁺	1	3 d	1.817	1.348	12		Fe ¹⁺	6	3 p	6.8	2.60	Est.
	Sc ³⁺	6	3 p	5.33	2.31	Est.			2	3 s	9.18	3.03	Est.
2	Ti	2	4 s	0.503	0.709	13, 18	27	Co	1	4 s	1.22	1.10	18
		2	3 d	0.730	0.855	13			2	4 s	0.577	0.759	18
		6	3 p	4.0	2.00	Est.		Co ¹⁺	1	4 s	1.27	1.126	18
		2	3 s	5.9	2.42	Est.			2	4 s	0.564	0.751	18
	Ti ¹⁺	1	4 s	1.003	1.001	13, 18	28	Ni	1	4 s	1.34	1.158	18
		2	3 d	1.2	1.10	Est.			1	4 s	1.34	1.158	18
	Ti ²⁺	2	3 d	2.038	1.428	13	29	Cu	1	4 s	0.567	0.754	1, 18
		1	3 d	3.301	1.817	14			10	3 d	0.773	0.881	3
	Ti ³⁺	6	3 p	7.07	2.66	Est.		Cu ¹⁺	10	3 d	1.503	1.225	18
3	V	2	4 s†	0.524	0.724	13	30	Zn	2	4 s	0.690	0.831	1, 18
		3	3 d	0.81	0.90	Est.			1	4 s	1.320	1.148	18
		6	3 p	4.1	2.15	Est.	31	Ga	1	4 p	0.440	0.664	1
		2	3 s	6.6	2.57	Est.			2	4 s	0.84	0.92	Est.
	V ¹⁺	4	3 d	1.041	1.020	13		Ge	2	4 p	0.70	0.84	Est.
	V ⁴⁺	1	3 d	5.072	2.252	14			2	4 s	1.00	1.00	Est.
	V ⁵⁺	6	3 p	9.0	3.00	Est.	33	As	3	4 p	0.86	0.929	1
									2	4 s	1.36	1.17	Est.
4	Cr	1	4 s	0.496	0.704	1, 18	34	Se	4	4 p	0.70	0.84	Est.
		5	3 d	0.610	0.781	3			2	4 s	1.40	1.18	Est.
		6	3 p	5.3	2.30	Est.		Br	5	4 p	0.86	0.93	Est.
		2	3 s	7.4	2.72	Est.			2	4 s	1.70	1.30	Est.
	Cr ¹⁺	5	3 d	1.23	1.11	18	36	Kr	6	4 p	0.98	0.991	19
									2	4 s	2.00	1.41	Est.
5	Mn	2	4 s	0.547	0.740	1, 18							
		5	3 d	0.67	0.82	Est.							
		6	3 p	6.0	2.45	Est.							
		2	3 s	8.24	2.87	Est.							
	Mn ¹⁺	1	4 s	1.161	1.076	18							
		5	3 d	1.25	1.12	Est.							

† See footnote 4.

that must be imparted to the atom or ion in its normal or most stable configuration, in order to remove that electron, and leave the resulting ion in its most stable state.⁴

The energy is expressed as a ratio of the energy of removal of the specified electron to that of the electron of the neutral hydrogen atom (for which the energy is 2.156×10^{-11} ergs, corresponding to an ionization potential of 13.54 volts). The energy values have been obtained in most cases by dividing the wave-number of the spectral term by the Rydberg wave-number 109678 cm.^{-1} . An attempt has been made to indicate the accuracy of the value by giving one more figure than is probably certain. In the last column of the table are shown the sources of the energy values, the numbers there corresponding to the references given at the end of the table. The letters Est. show that the energy value is estimated and is not the result of direct determination. These estimated values have for the most part been derived with the aid of the principle established by Millikan and Bowen (2) that, for atoms and ions having the same electronic configuration and differing only in their nuclear charge (for example, Ne, Na⁺, Mg²⁺, Al³⁺, Si⁴⁺), the increase in the square-root of the energy of removal of an electron from any specified subshell is nearly proportional to the increase in the nuclear or atomic number.

PERIODIC CHART SHOWING THE ENERGIES OF REMOVAL AND THE QUANTUM STATES OF THE ELECTRONS IN THE OUTER SHELLS OF ATOMS

With the aid of the data presented in table 1, the chart shown on the following pages has been prepared. The conventions employed are as follows.

Along the axis of abscissas is plotted the atomic number of the element, which also represents the positive charge on the

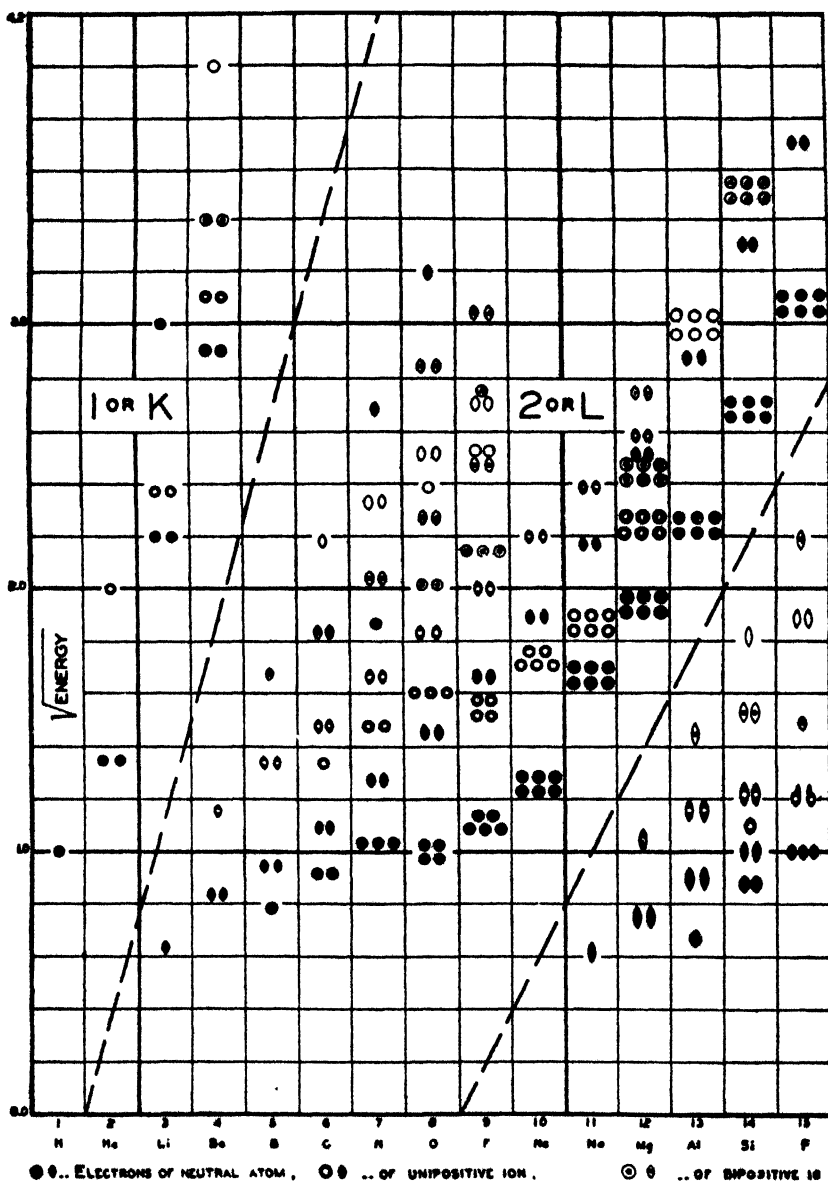
⁴ In general, there is no change in the quantum states of the remaining electrons; but to this there are occasional exceptions. Thus, when the neutral vanadium atom loses one of its two 4s electrons, the other 4s electron changes to a 3d state, associating itself with the three electrons already in this state. (In spectroscopic notation this is expressed by saying that the atom configuration changes from 4s²3d³ to 3d⁴, and not to 4s 3d³; the superscripts representing the number of electrons existing in the state to whose symbol they are attached.)

nucleus of its atom and ions (also the total number of electrons around the nucleus of the neutral atom). As ordinates are plotted the square-roots of the energies absorbed in the removal of the respective electrons from the atom, referred to that absorbed in removing the electron from the hydrogen atom as unity; in other words, the square-roots of the ratios ν / R of the wave-number of the spectral term to the Rydberg number. The square-roots are plotted, rather than the energies themselves, since this enables higher energy values to be included in the chart without unduly extending it, and since Millikan and Bowen (2) found that the square-roots of the energies of removal of the outer electrons of atoms and ions that have the same electronic configuration but different nuclear charges, are nearly linear functions of their atomic numbers.

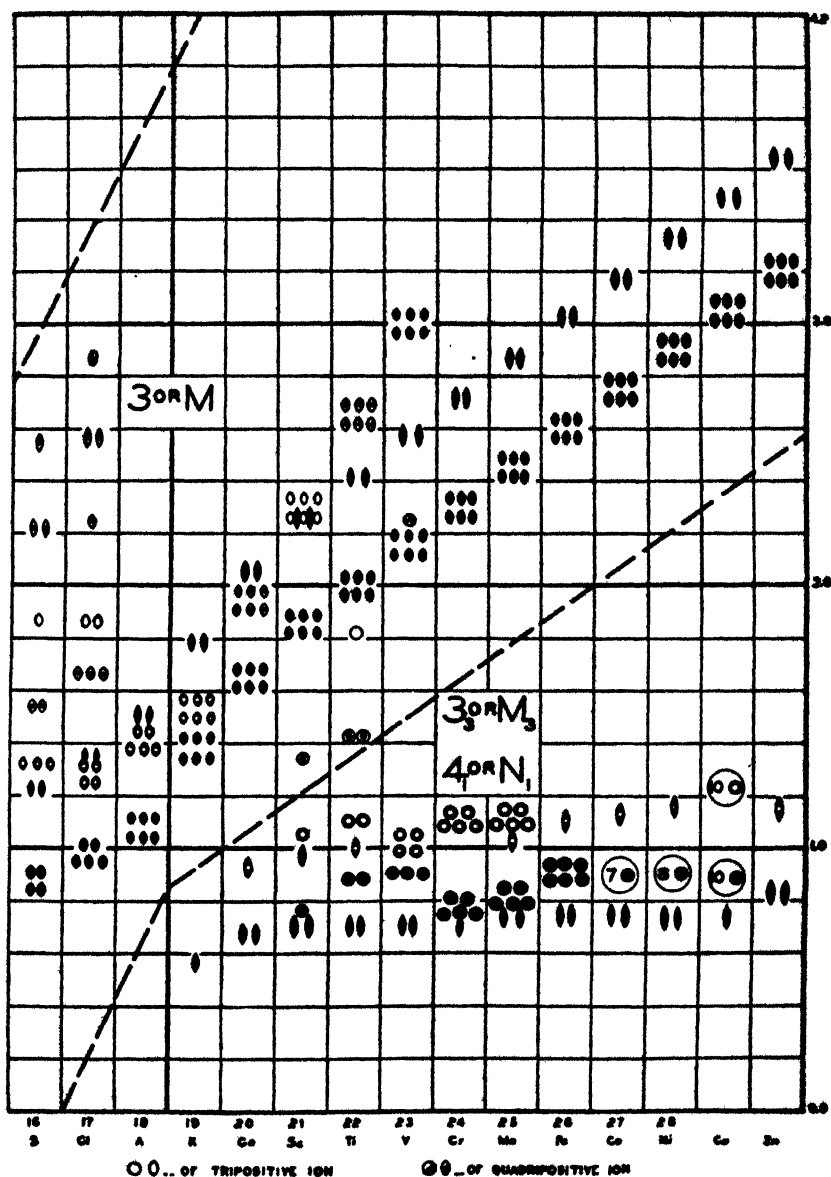
The number of electrons having a specified energy of removal is shown by a corresponding number of circles or ellipses; it being understood that the exact value of the energy is at the center around which these electrons are grouped. Whether the electron is removed from the neutral atom of an element or from one of its successive ions is indicated by different types of circles or ellipses, as shown by the legend at the bottom of the chart.*

The quantum-states of the electrons are differentiated by the use of circles and of ellipses of different eccentricity. Namely, when the total and azimuthal quantum numbers (n and k) are equal (as in the 1_1 , 2_2 , or 3_3 quantum states), the electrons are represented by circles. When these quantum numbers are unequal, ellipses are used whose major and minor axes have the same ratio as the two quantum numbers, thus 2:1 for a 2_1 state; 3:2 for a 3_2 state; etc. The total quantum numbers (1, 2, 3, 4) or the shells (K, L, M, N) of the electrons are also indicated by the regions of the plot in which they lie, these regions being separated from each other by heavy oblique broken lines. It was not possible, however, to separate completely in this way the

* In making a wall chart for lecture demonstration the authors have used, in order to make more prominent the distinction between the atom and its ions, in place of these different symbols, different colors (black, red, orange, green, blue) for the neutral atom and the successive positive ions.



THE STRUCTURE OF ATOMS



THE STRUCTURE OF ATOMS

3 or M from the 4 or N electrons; for the electrons of the 3, and 4, quantum types have energies of removal of the same order of magnitude. In the lower right-hand region, therefore, the circles are to be understood to represent 3, electrons and the ellipses 4, electrons.*

The chart gives a complete representation of all the electrons for the neutral atom and all its ions only in the case of the first four elements (hydrogen to beryllium). For all the other elements the two electrons in the K shell are omitted, as they would lie far above the top of the chart; and for the neutral atoms above phosphorus and for the ions above magnesium the eight electrons of the L shell are also omitted. These omissions must be borne in mind in considering the complete structure of the atom. No attempt is made to indicate on the chart whether the points are based on estimated or directly measured data; but this was shown in the table of data in the preceding section.

THE GROUPS OF THE PERIODIC SYSTEM IN RELATION TO THE STRUCTURE OF ATOMS

Now that we have more complete knowledge of the energies of removal of the electrons as well as of their quantum states, it seems desirable to review the general relations between the structure of atoms and the periodicity of the properties of the elements.

Attention may first be called to the extent of the similarities

* These conventions may be illustrated by reference to the element beryllium. It is seen that the neutral atom of beryllium has two (loosely-bound) electrons of the 2, quantum type (in the L shell) with an energy of removal equal to $(0.838)^2$ or about 0.70 that of the hydrogen electron, and two (much more firmly bound) electrons of the 1, quantum type (in the K shell) with an energy of removal equal to $(2.9)^2$ or 8.5 times that of the hydrogen electron. The unipositive ion Be^+ is seen to have its single remaining 2, electron at a higher level than before, with an energy of removal $(1.157)^2$ or 1.34 times that of the hydrogen electron, and it has its two 1, or K electrons at a level corresponding to an energy of removal equal to $(3.1)^2$ or 9.6 times that of the hydrogen electron. The chart also shows that the bipositive ion Be^{++} has only the two 1, or K electrons remaining, that these have an energy of removal equal to $(3.38)^2$ or 11.4 times that of the hydrogen electron. Finally, the tripositive ion Be^{+++} is seen to have only one K electron, which has an energy of removal $(4.00)^2$ or 16.0 times that of the hydrogen electron.

exhibited not only by the quantum states, but also by the energies of removal of the outer-shell electrons of the neutral atoms of the elements of the same group in the first three periods of the periodic system. To make these relations clearer, there have been brought together in table 2 the numbers, quantum states, and energies of removal of the electrons for the neutral atoms of the first three elements of each of the eight periodic groups, as well as for the element of the other subgroup in the third period of each group. These values have been taken from table 1.

Examination of table 2 shows that the atoms of the first three elements in each group have the same number of electrons in the same azimuthal quantum state, and have energies of removal of the same general magnitude, there being however usually a considerable decrease in the energy in passing from the first to the second and third elements. On the other hand, comparison of the first three elements of groups III to VII with the fourth element (which belongs to the other subgroup of the periodic system) shows that the azimuthal quantum states of the outer electrons of the atom are not the same—facts which are no doubt, related to the marked difference in properties. This difference in the azimuthal quantum states is not shown in groups I and II, so that the difference in properties of the alkali elements and copper, or of the alkaline-earth elements and zinc, must be due to other factors.

The atomic structure also accounts for the fact that the third period includes eighteen elements, instead of eight. Referring to the chart and comparing the second and third periods, the following differences are noted. In the second period, in passing from magnesium 12 to argon 18, the two $3s$ electrons steadily become more and more firmly bound; and when other electrons are added, as in the elements from aluminum 13 to argon 18, these are of the $3p$ type and have progressively increasing⁷ energies of removal much less than those of the $3s$ electrons, a condition which accounts for the progression in physical properties. In the third period, on the other hand, though it starts

⁷ Except that there is a break in the progression at sulfur 18.

TABLE 2

The outer-shell electrons of corresponding elements of different periodic groups

GROUP	ATOM	OUTER ELECTRONS			GROUP	ATOM	OUTER ELECTRONS		
		Number	State	Energy			Number	State	Energy
I	Li	One	2 ₁	0.40	V	N	Three	2 ₂	1.07
	Na	One	3 ₁	0.38			Two	2 ₁	1.62
	K	One	4 ₁	0.32		P	Three	3 ₂	0.98
	Cu	One	4 ₁	0.57			Two	3 ₁	1.52
II	Be	Two	2 ₁	0.70	VI	As	Three	4 ₂	0.86
	Mg	Two	3 ₁	0.56			Two	4 ₁	1.36
	Ca	Two	4 ₁	0.45		V	Two	4 ₁	0.52
	Zn	Two	4 ₁	0.69			Three	3 ₂	0.81
III	B	One	2 ₂	0.62	VII	O	Four	2 ₂	1.00
		Two	2 ₁	0.89			Two	2 ₁	2.09
	Al	One	3 ₂	0.44		S	Four	3 ₂	0.76
		Two	3 ₁	0.81			Two	3 ₁	1.49
IV	Ga	One	4 ₂	0.44	VIII	Se	Four	4 ₂	0.70
		Two	4 ₁	0.84			Two	4 ₁	1.40
	Sc	Two	4 ₁	0.49		Cr	One	4 ₁	0.50
		One	3 ₂	0.55			Five	3 ₂	0.61
	C	Two	2 ₂	0.84	IX	F	Five	2 ₂	1.25
		Two	2 ₁	1.18			Two	2 ₁	2.74
	Si	Two	3 ₂	0.76		Cl	Five	3 ₂	0.96
		Two	3 ₁	1.00			Two	3 ₁	1.81
	Ge	Two	4 ₂	0.70	X	Br	Five	4 ₂	0.86
		Two	4 ₁	1.00			Two	4 ₁	1.70
	Ti	Two	4 ₁	0.50		Mn	Two	4 ₁	0.55
		Two	3 ₂	0.73			Five	3 ₂	0.67
					XI	Ne	Six	2 ₂	1.59
							Two	2 ₁	3.57
						A	Six	3 ₂	1.16
							Two	3 ₁	2.25
					XII	Kr	Six	4 ₂	0.98
							Two	4 ₁	2.00

normally by adding 4_1 electrons, the conditions are complicated by the fact that there exists in the atoms an unfilled 3 shell, namely the 3_s subshell. From this subshell the energy of removal, though nearly the same as from the 4_1 subshell, is much greater than from the 4_s subshell. Consequently, the electrons next added are of the 3_s type, not of the 4_s type which would correspond to the 3_s type of the second period. An additional fact is that the energies of the 3_s and 4_1 electrons retain nearly the same values for the successive elements until the 3_s subshell is completely filled with its ten electrons. This results in a long series of similar elementary substances which have atoms with two electrons of the 4_1 type and a group of 3_s electrons, both types having nearly the same energies of removal.⁸ And it is only beyond zinc (when the 4_1 and 3_s subshells have been filled) that the new electrons begin to take the six places existing in the 4_s subshell, thereby giving rise to six elements, gallium 31 to krypton 36 (not shown in the chart) with properties very similar to those of the last six elements of the second period (aluminum 13 to argon 18) which are characterized by loosely bound 3_s electrons, thus with the same azimuthal quantum number.

It is perhaps well to emphasize the fact, pointed out early by investigators of atomic structure, that periodic tables of the form used by J. Thomsen, Bohr, and others⁹ in which the first and second periods consist of eight elements and the third period of eighteen elements, and the former periods are shown to be in a dual relationship to the latter, are a much better expression of atomic-structure results than the more commonly used tables, in which, as in that of Lothar Meyer, the elements are so far as possible classified in successive periods of eight.

⁸ There is, to be sure, a progressive increase in the number of 3_s electrons which makes possible compounds of higher valences, increasing progressively from scandium to iron.

⁹ Charts of this type which show the relationships rather strikingly have recently been described by von Antropoff (*Z. angew. Chem.*, **39**, 722 (1926)), and are sold by Köhler and Volckmar of Leipzig, Germany.

ION-FORMATION AND VALENCE IN RELATION TO THE STRUCTURE OF ATOMS

With the aid of the chart and the data of tables 1 and 2 there may now be studied empirically¹⁰ the relations that exist between the structures of the atoms of the elements, and (a) the nature of the ions which the elements form in crystals and aqueous solutions, and (b) the valences which the elements actually exhibit in their compounds.

The characteristics of the two classes of compounds, known as polar and non-polar, are fairly distinct, even though they doubtless grade into each other. In polar compounds, of which salts in the crystalline, fused, and dissolved states are examples, there exist, more or less separate from each other, positive and negative ions formed from neutral atoms by the loss or gain of one or more electrons. In non-polar compounds the electronic condition of the united atoms has been the subject of many hypotheses. For our purposes it will, however, suffice to point out that the valence shown by an element in its non-polar compounds doubtless corresponds to the existence of a certain number of electrons in such a state that they take part in the process of formation of the molecules of the compounds. The electrons in this state may be called the *reacting electrons* for that valence.

To facilitate the study of the relation of these characteristics

¹⁰ There cannot be discussed here the highly important theoretical relations that have been recently developed. Reference may, however, be made to the original work of Grimm and Herzfeld (Z. Physik., **19**, 141 (1923)), according to which only those kinds of ions exist in crystals or solutions whose formation from the neutral elementary substances is attended by a net decrease in energy-content as computed from the energies of removal of the electrons of the neutral atoms and from the other energy effects involved; namely, from the energies of vaporization of the metal and of dissociation of the diatomic gases (Cl₂, O₂, etc.), and from the "lattice energy" attending the bringing together in the crystal of the gaseous ions, or the electrostatic energy attending their introduction into a dielectric medium such as water. The latter subject has been also treated by Webb (J. Am. Chem. Soc., **48**, 2589 (1926)). Of transcendent importance to our interpretation of valence and non-polar bonding may also become the theories based on wave mechanics of London (Z. Physik., **46**, 455 (1928)), which confirm the earlier conclusion of Lewis that single chemical bonds consist in the pairing of two electrons of two different atoms. (See also Pauling, Proc. Nat. Acad. Sci., **14**, 359 (1928).)

TABLE 3

Ion-formation and valence in relation to atomic structure

ELEMENT	ATOMIC NUMBER	IONS FORMED	VALENCES	IONS OR ATOMS WITH		ENERGY OF REMOVAL OF OUTER ELECTRONS FROM				
				Complete shell	Complete subshell	M	M ⁺	M ²⁺	M ³⁺	M ⁴⁺
H	1	H ⁺ H ⁻	1 -1	H ⁺ H ⁻	..	1.00	∞
He	2	None	0	He	..	1.80	4.00	∞
Li	3	Li ⁺	1	Li ⁺	..	0.40	5.6	9.00	∞	...
Be	4	Be ²⁺	2	Be ²⁺	..	0.70	1.34	11.4	16.0	∞
B	5	...	3	B ⁺	B ⁺	0.62	1.79	2.80	19.3	25.0
C	6	...	4	C ⁴⁺	C ²⁺	0.83	1.79	3.36	4.74	29.2
N	7	...	3, 5 -3	N ⁵⁺ N ³⁻	N ²⁺	1.07	2.18	3.49	5.43	7.19
O	8	O ⁶⁺	O ⁴⁺	1.00	2.58	4.05	5.69	8.07
F	9	O ²⁻ ...	-2 ...	O ²⁻ F ⁷⁺	F ⁶⁺	1.25	2.39	4.62	6.40	7.56
Ne	10	F ⁻ None	-1 0	F ⁻ Ne	..	1.59	3.02
Na	11	Na ⁺	1	Na ⁺	..	0.38	3.47
Mg	12	Mg ²⁺	2	Mg ²⁺	..	0.56	1.11	5.98
Al	13	Al ³⁺	3	Al ³⁺	Al ⁺	0.44	1.34	2.09	9.00	...
Si	14	...	4	Si ⁴⁺	Si ²⁺	0.76	1.20	2.34	3.32	12.47
P	15	...	1, 3, 5 -3	P ⁵⁺ P ³⁻	P ²⁺	0.98	1.46	2.22	3.53	4.78
S	16	..	4, 6 -2	S ⁶⁺ S ²⁻	S ⁴⁺	0.76	1.74	2.37	3.48	4.98
Cl	17	...	1, 3, 5, 7	Cl ⁷⁺	Cl ⁵⁺	0.96	1.66	2.93	3.50	5.00
A	18	Cl ⁻ None	-1 0	Cl ⁻ A	..	1.16	2.00
K	19	K ⁺	1	K ⁺	..	0.32	2.34
Ca	20	Ca ²⁺	2	Ca ²⁺	..	0.45	0.87	3.76
Sc	21	Sc ³⁺	3	Sc ³⁺	..	0.49	0.95	1.82	5.33	...
Ti	22	Ti ³⁺ Ti ²⁺ Ti ⁴⁺	2, 3, 4	Ti ⁴⁺	..	0.50	1.00	2.04	3.30	7.07
V	23	V ³⁺ V ²⁺	2, 3, 4 5	V ⁵⁺	..	0.52	1.04	5.07
Cr	24	Cr ²⁺ Cr ³⁺	2, 3, 6	Cr ⁶⁺	..	0.50	1.23
Mn	25	Mn ²⁺ Mn ³⁺	2, 3, 4, 6, 7	Mn ⁷⁺	..	0.55	1.16
Fe	26	Fe ²⁺ Fe ³⁺	2, 3, 6	Fe ⁶⁺	..	0.58	1.22

to ion-formation and valence, table 3 has been prepared from the atomic-structure data given in tables 1 and 2 and the known chemical properties of the elements. The headings are for the most part self-explanatory; but the following supplementary statements may be made. The "ions formed" are all the simple elementary ions of whose actual existence in aqueous solutions or in crystals there is good evidence. The indicated "valences" include only those corresponding to the well-defined states of oxidation that are exhibited by the elements in their salts, bases, or acids. The symbols of the "ions or atoms" in the next two columns show those configurations which have, or would have if they existed, complete shells or complete subshells. Note that the charges on these hypothetical ions show also the *valences* that would exist if the lost or gained electrons were actually the reacting electrons in the neutral atom.

A study of the table leads to the following conclusions.

1. In the case of the first three elements of the first two periods (except boron) positive ions are formed at the point where the electrons remaining just suffice to complete the next lower main shell (the 1 or 2 shell). This point is, moreover, identical with the point where the removal of a further electron would require a much greater absorption of energy. Thus, the ions formed are Li^+ , Be^{2+} , Na^+ , Mg^{2+} , and Al^{3+} , all of which are "stripped atoms" so-called. These elements do not form ions at points intermediate between the stripped atom and the neutral atom, even in cases where the energies of removal of the successive electrons differ by considerable amounts, or in cases where it would leave as the outermost electrons a complete subshell (2_1 or 3_1). Thus the ions Be^+ , Mg^+ , and Al^+ do not form in solutions or crystals, even though the energy of removal of the second electron is two or three times that of the first electron; and Al^{2+} does not form, even though the removal of its 3_2 electron would leave its 3_1 shell complete.

2. The first three elements of the third period, potassium, calcium, and scandium, form through the loss of their 4_1 electrons (and of the 3_2 electron of scandium) the ions K^+ , Ca^{2+} and Sc^{3+} ; even though there then remains not a complete main shell (the 3 shell), but only two fully filled subshells (3_1 and 3_2).

3. Each of the first three elements of the first three periods exhibits in its chemical compounds only that valence which is identical with the charge on the ion which it actually forms; showing that the readily detachable electrons in the neutral atom are also those concerned in the formation of molecules.

4. The middle elements, carbon and silicon, of the first two periods do not seem to form ions; but they show the valence +4, in correspondence with the tendency of all the electrons of a neutral atom that are outside of a completed shell to be reactive.

5. The elements near the ends of the first and second periods do not form positive ions, but produce such negative ions as correspond to the completion of the partially filled shell; thus these elements form the ions O^- , F^- , S^- , Cl^- .

6. Nitrogen in the first period, and phosphorus, sulfur, and chlorine in the second period, all exhibit very stable (maximum) positive valences that correspond to the tendency of all the electrons that are outside of a complete shell to be reactive. These elements also exhibit the negative valences that correspond to that number of electrons which would complete the next main shell. Thus nitrogen or phosphorus has the valence +5 in nitrates or phosphates, and the valence -3 in ammonia or phosphine; sulfur has the valences +6 in sulfates and -2 in sulfides; and chlorine has the valences +7 in perchlorates and -1 in chlorides.

These elements also exhibit other positive valences, which may be considered in relation to the various subshells. Thus in the case of nitrogen or phosphorus the electrons present in the neutral atom in excess of those needed to complete the 2_1 or 3_1 subshell correspond to the valence +3; and this valence is in fact shown in nitrites or phosphites. In the case of sulfur the excess electrons outside the 3_1 subshell correspond to the valence +4, which exists in sulfites. And in the case of chlorine the excess electrons outside the 3_1 subshell correspond to the valence +5, which exists in chlorates.

These elements may exhibit, however, besides these valences related to complete shells or subshells, other less stable valences. These valences correspond to a diminished reactivity of the

successive pairs of electrons remaining after the most reactive electrons of a given subshell have entered into molecular unions. Thus phosphorus shows the valence +1 in the unstable hypophosphites, indicating that of the three electrons in the $3s$ subshell of its atom two of them are less reactive, at any rate after one of them has entered into chemical union. Similarly, sulfur has the valence +2 in derivatives of the non-existent H_2SO_2 , and also perhaps in the unstable hydrosulfurous acid $H_2S_2O_4$ (if this be regarded as an anhydride of H_2SO_2 and H_2SO_3); and chlorine shows the valence +3 and +1 in the unstable chlorites and hypochlorites. These valences of sulfur and chlorine correspond to the fact that of the electrons of the $3s$ subshell one pair (or two pairs) are less reactive.

7. The elements of the third period from titanium 22 to copper 29 form a series of elementary ions of varying charge, and exist in a series of positive valences which are not merely even or odd as in the case of the elements of the first two periods, but increase continuously or irregularly. Thus titanium forms the ions Ti^{2+} , Ti^{3+} , and Ti^{4+} , and exists in corresponding valences; vanadium forms V^{3+} and V^{5+} , and exists not only in the corresponding valences, but also in the +4 and +5 stages; chromium forms Cr^{2+} and Cr^{3+} , and shows the valences +2, +3, and +6, but not +4 and +5; manganese forms Mn^{2+} and Mn^{3+} , and shows also the valences +4, +6, and +7, but not +5; iron forms Fe^{2+} and Fe^{3+} , and has the valences +2, +3, and +6 (in ferrates). Unfortunately, energy data are still lacking for many of the ions of these elements (except in the case of titanium); but the results lead to the following conclusions.

The maximum valence of the elements from titanium to manganese corresponds to the participation of all the $4s$ and $3d$ electrons in chemical combinations, leaving the unreacting electrons in a complete $3s$ subshell. The formation by all the elements from titanium to zinc of stable bipoisitive ions corresponds to the loss of the two $4s$ electrons (except in the cases of chromium and copper); and the bipoisitive valence of these three elements corresponds to the participation of these two $4s$ electrons in chemical unions. The electrons in the $3s$ subshell are evidently

less reactive, even though this subshell is not completely filled, indicating that an outermost 3, subshell (unlike the lower quantum subshells), even when not completed, constitutes an unusually stable configuration, moreover, one whose stability seems to be little affected by the number of electrons which it contains. Beyond this, however, our present knowledge of atomic structure does not go in interpreting the peculiar valence relations of these elements. It seems clear, however, as pointed out by G. N. Lewis (29) that the irregular valences of these metal-forming elements constitute a phenomenon of a different type from the even or odd valences exhibited by the elements of the nitrogen, sulfur, and chlorine groups.

SUMMARY

In this article has been presented a compilation of values, derived from recent spectroscopic studies, of the energies attending the removal of the separate electrons from the atoms and ions of the first thirty elements. With the aid of these values and data as to quantum states, a periodic chart has been prepared showing graphically the number, the quantum states, and the energies of removal of the electrons in the outer shells of these atoms and ions.

These atomic structures are discussed in relation to the periods and groups of the periodic system of the elements. The structures are also discussed in detail in relation to the formation of ions in crystals or solutions by the separate elements, and in relation to the best-defined valences which they exhibit.

The conclusions as to ion-formation may be briefly summarized as follows:

Of the elements preceding titanium, the first three elements of the first three periods form only such positive ions as result from the removal of all the electrons outside of the complete 1, 2, or $3_1 + 3_2$ shells; and the last three elements of each period form only such negative ions as result from the complete filling of the 2 or $3_1 + 3_2$ shells with electrons. Intermediate ions, such as Al^+ or Ca^+ , do not form, even when, as with Al^+ , the resulting ion would

have its outermost electrons in a complete subshell (3_1). Titanium, vanadium, chromium, manganese, iron, cobalt, nickel, and copper all form bipoisitive ions, even though there remain electrons in an unfilled 3_s subshell. They also form other ions which have no apparent relation to the structure of their atoms.

The conclusions as to valence may be summarized as follows:

None of the first three elements of the first three periods exhibit any well-defined valence other than that which would result if the electrons giving rise to the valence were those, and only those, outside of the complete 1, 2, or 3_1+3_s shells. The possibility that, in the cases of boron and aluminum or of carbon and silicon, the electrons outside of a complete subshell (2_1 or 3_1) might alone function as reacting electrons does not cause these elements to actually exist in the univalent or bivalent state.

The elements nitrogen, phosphorus, sulfur, and chlorine, on the other hand, exhibit the positive valences that would result if the reacting electrons were:

- (1) All those in an unfilled main shell (2 or 3) outside a filled main shell (1 or 2).
- (2) All those in an unfilled subshell (2_s or 3_s) outside a filled subshell (2_1 or 3_1).
- (3) Those in an unfilled subshell (2_s or 3_s) in excess of one pair or two pairs of (less reactive) electrons in the same subshell.

Thus the positive valences that would correspond to these three types are:

- (1) Nitrogen, +5; phosphorus, +5; sulfur, +6; chlorine, +7.
- (2) Nitrogen, +3; phosphorus, +3; sulfur, +4; chlorine, +5.
- (3) Nitrogen, +1; phosphorus, +1; sulfur, +2; chlorine, +3.

And in fact these elements exist in the form of salts in these valences. Furthermore, the chemical stability of these various valences is clearly related to the type, being greatest for type (1) and least for type (3). Thus, type (1) is that of the very stable nitrates, phosphates, sulfates, and perchlorates; type (2) is that of the much less stable nitrites, phosphites, sulfites, and chlorates; and type (3) is that of the unstable hypophosphites, chlorites, and hypochlorites, and of the extremely unstable or even non-existent salts MNO and M_2SO_3 .

The first eight elements of the third period exhibit (maximum)

positive valences that correspond to a reactivity of the electrons outside of the completed 3, or 3, subshells. Moreover, all of them beyond scandium form bivalent compounds, which indicates that the 3, subshell is an exceptionally stable configuration, even when it is incompletely filled with electrons. The other irregular valences exhibited by these elements are not related in any apparent way to the structure of their atoms.

REFERENCES FOR TABLE 1

- (1) RUSSELL: *Astrophys. J.*, **61**, 223 (1925).
- (2) MILLIKAN AND BOWEN: *Proc. Nat. Acad. Sci.*, **13**, 531 (1927).
- (3) BOWEN: Preliminary unpublished value.
- (4) BOWEN AND MILLIKAN: *Phys. Rev.*, **27**, 144 (1926).
- (5) HOLWECK: *Compt. rend.*, **180**, 658 (1925).
- (6) FOWLER: *Phil. Trans. Roy. Soc. Lond.*, **A 225**, 1 (1925).
- (7) BOWEN: *Phys. Rev.*, **29**, 510 (1927).
- (8) BOWEN AND MILLIKAN: *Phys. Rev.*, **25**, 600 (1925).
- (9) BOWEN AND MILLIKAN: *Phys. Rev.*, **25**, 295 (1925).
- (10) SAUNDERS: *Proc. Nat. Acad. Sci.*, **12**, 556 (1926).
- (11) RUSSELL AND MEGGERS: *Phys. Rev.*, **29**, 606 (1927).
- (12) SMITH: *Proc. Nat. Acad. Sci.*, **13**, 65 (1927).
- (13) RUSSELL, DUGAN, AND STEWART: *Astronomy*, Ginn and Co. 1927, Table 38.
- (14) GIBBS AND WHITE: *Proc. Nat. Acad. Sci.*, **12**, 598 (1926).
- (15) RUSSELL, COMPTON, AND BOYCE: *Bull. Amer. Phys. Soc.*, **3**, 15 (1928).
- (16) BOWEN: *Phys. Rev.*, **31**, 34 (1928).
- (17) BOWEN: *Phys. Rev.*, **31**, 497 (1928).
- (18) RUSSELL: *Astrophys. J.*, **66**, 233 (1927).

OTHER REFERENCES

- (19) MAIN-SMITH, *Chemistry and Atomic Structure*, D. Van Nostrand Co., N. Y., 1924, p. 196-198.
- (20) STONER: *Phil. Mag.*, **48**, 719 (1924).
- (21) MCLENNAN, MCLAY, AND SMITH: *Proc. Roy. Soc. London*, (A) **112**, 76 (1926).
- (22) ANDRADE: *The Structure of the Atom*, Harcourt, Brace and Co., N. Y., 1927, p. 724-727.
- (23) BOHR: *Nature*, **112**, 29 (1923).
- (24) MAIN-SMITH: *Chemistry and Atomic Structure*, p. 189-198.
- (25) RUSSELL AND SAUNDERS: *Astrophys. J.*, **61**, 38 (1925).
- (26) PAULI: *Z. Physik.*, **31**, 765 (1925).
- (27) HEISENBERG: *Z. Physik.*, **32**, 841 (1925).
- (28) HUND: *Z. Physik.*, **33**, 345 (1925); **34**, 296 (1925).
- (29) LEWIS: *Valence and the Structure of Atoms and Molecules*, Chemical Society Monograph, 1923, p. 61.

LINE SPECTRA AND THE PERIODIC ARRANGEMENT OF THE ELEMENTS

SAUL DUSHMAN

Research Laboratory, General Electric Company, Schenectady, N. Y.

INTRODUCTION

The conception of the periodic system is the most important generalization in the whole domain of chemistry. It is recognized now, as a result of the work of Rutherford, Moseley, Aston, and others, that the properties of the elements are periodic functions, not of the atomic weight (as was believed for so many years) but of the atomic number (N). Since the atomic number corresponds to the charge on the nucleus, and since the number of electrons external to the nucleus must be equal to the nuclear charge, it would seem to be a logical consequence to correlate the periodic properties of the elements with periodic variations in the distribution of electrons in the various atoms.

Such a correspondence has actually been derived, as a result of the Bohr theory, from the study of atomic spectra, both in the visible and x-ray region. While the work of the early spectroscopists had led chemists to use spectra merely as a qualitative guide towards the identification of elements, and physicists had previously found it impossible to formulate an atomic model by which to account for the origin of spectra, Bohr's conception of energy "levels" and electronic orbits has paved the way for a well-nigh complete revelation of the structure of the atom in all its complexity. It is true that there are still a great many problems regarding atomic structure which remain unanswered, the whole conception of electronic orbits is being viewed askance by a large and influential group of physicists, and the wave mechanics bids fair to outlaw the "naïve" idea (as one physicist has designated it) of orbits. But nevertheless, the most essential concept in Bohr's theory—that of the correspondence between so-called

energy levels in the atom and spectral "terms"—remains unaltered and provides the guiding principle in correlating the periodic properties of the elements with their spectra.

BOHR THEORY, ENERGY LEVELS

In order to understand the method of interpreting spectra in terms of the arrangement of electrons in atoms it is necessary to recapitulate briefly the essential features of the Bohr theory.

The assumption on which Bohr bases his theory of the origin of spectra has been stated by him in the following form (1).

Postulate A.

An atomic system can, and can only, exist permanently in a certain series of states corresponding to a discontinuous series of values for its energy, and consequently any change of the energy of the system, including emission and absorption of electromagnetic radiation, must take place by a complete transition between two such states. These states will be denoted as the "stationary states" of the system.

Postulate B. That the radiation absorbed or emitted during a transition between two stationary states is *monochromatic* and possesses a frequency ν , given by the relation

$$h\nu = E_i - E_f \quad (1)$$

where E_i and E_f designate the energy values corresponding to initial and final states respectively, and h is Planck's constant.

In order to determine the individual energy values of the different stationary states Bohr assumed that the angular momentum of the electron in its orbit is an integral multiple of $h/2\pi$. It was shown later by Wilson and Sommerfeld that this assumption could be replaced by a more generalized one which is based on classical mechanics and states that for each degree of freedom of an atomic system

$$\oint p dq = n h \quad (2)$$

where q is a generalized Lagrangian coordinate, and p is the corresponding momentum, and the integral is taken over a

complete period of the motion. The integer n is known as the quantum number for the particular degree of freedom concerned.

Representing the hydrogen atom as constituted of an electron rotating in a Keplerian orbit around a nucleus of unit positive charge, Bohr derived the following relation for the energy of the atom in any state of quantum number n .

$$E_n = -Rch/n^2 \quad (3)$$

where c = velocity of light and R = a universal constant.

Hence, in accordance with equation (1), the frequency of any line in the spectrum of atomic hydrogen should be represented by a relation of the form

$$\nu = Rc \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \quad (4)$$

where m and n are the quantum numbers of initial and final states respectively.

Now this is exactly the type of equation which Balmer had found a long time ago to apply to a series of lines in the visible portion of the hydrogen spectrum, and slightly modified forms of this equation (with the same constant R) had been used by Ritz and Rydberg to represent certain series observed in the spectra of the alkalis, alkaline earths, and some of the other elements. Thus Bohr gave a theoretical interpretation of a conclusion which had been empirically derived by spectroscopists, that the frequency of a line is to be represented as a difference between two terms, one of which (that corresponding to n in the above equation) remains constant for all the members of the same series, while the other (corresponding to m) varies by unity for successive members of the series.

However, Bohr went one step further and actually found that from his model of the hydrogen atom the value of the Rydberg constant, R , could be calculated from fundamental constants by the relation

$$R = \frac{2\pi^2 e^4 m_0}{ch^3} = 109,737 \text{ cm}^{-1}$$

where e = charge, and m_0 = mass of electron.

Bohr also pointed out the manner in which the value of this constant must be modified slightly for different values of the atomic mass.

For a "hydrogen-like" atom consisting of a positive nucleus of charge Ze and an electron rotating about this nucleus, the energy corresponding to the orbit of quantum number n is given by

$$E_n = - \frac{RchZ^2}{n^2} \quad (5)$$

Physically, the expression Z^2Rch/n^2 is to be interpreted as the amount of energy required to remove the electron from the particular orbit to infinity. Since this energy may be determined experimentally, in the case of the normal atom, by bombarding the latter with electrons possessing a definite velocity, it is usually expressed in volts, (energy = Ve) and is known as the *ionizing potential* (V_i). Similarly the energy corresponding to the removal of the electron from its normal orbit to that of the first excited state is known as the *resonance potential* (V_r).

In the case of the hydrogen atom

$$V_i e = Rch/1^2$$

since $n = 1$, $m = \infty$.

Substituting the known values of R , c , h and e , it is found that $V_i = 13.54$ volts.

Similarly, $V_r e = Rch \left[\frac{1}{1^2} - \frac{1}{2^2} \right]$, and $V_r = 10.16$ volts.

In the case of other elements, these critical potentials (as they are known) have been determined either by direct measurement (2) or from the values of the lowest energy levels. The discussion of the data thus obtained is, however, best deferred for a subsequent section.

SPECTRAL SERIES. QUANTUM NUMBER, k

While in the case of the hydrogen atom the different spectral terms may be designated (to a very high degree of approximation) by integral values of a single quantum number, this is no longer

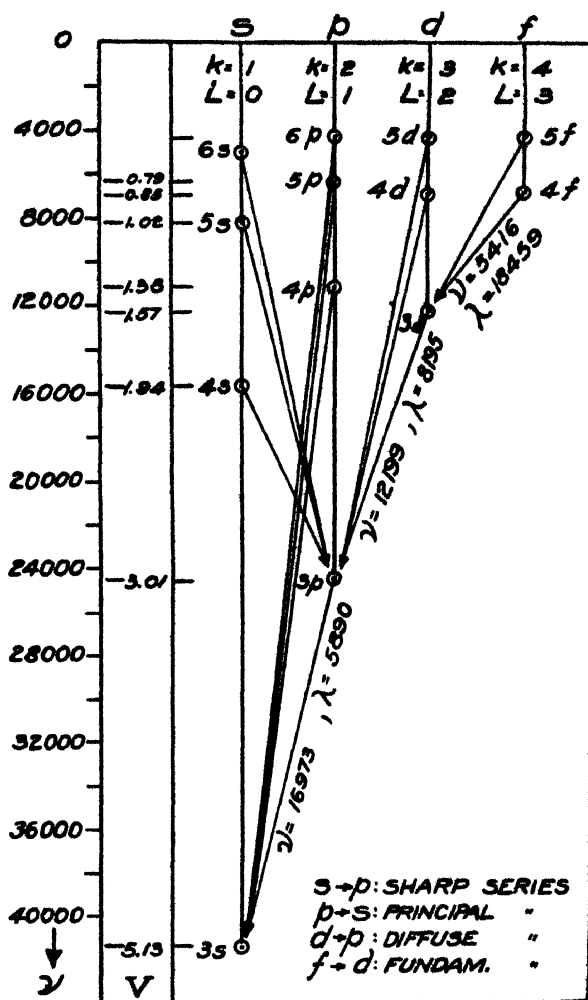


FIG. 1. ENERGY LEVELS FOR SODIUM

possible in the case of the other elements. In fact, in order to account for the spectral structures observed for the latter and also for the behavior of their lines in magnetic fields, it has been found necessary to introduce at least three other quantum numbers by which each term is defined. Before, however, discussing this, it is essential to describe some of the main features of spectroscopic nomenclature.

As a typical spectrum we may consider that of sodium. Here it is found that the observed lines group themselves into four different series and must, therefore, be regarded as due to transitions between four different sets of terms or corresponding energy levels in the atomic system. Figure 1 gives a diagrammatic representation of these levels. These four series of terms are designated by the letters *s*, *p*, *d* and *f* respectively. Now it will be observed that transitions occur only between neighboring series of levels, but *never* from one level in any one series to another level in the same series. This is due to the existence of a "selection principle" which forbids transitions between orbits for which the value of a second quantum number *k* is the same, and permits only those transitions to occur for which the value of *k* changes by +1 or -1. Physically *k* has been regarded as a measure of the angular momentum of the electron in its orbit in terms of multiples of $h/2\pi$ and is known as the azimuthal quantum number. Thus the level 3*s* has the total quantum number 3 and electronic angular momentum of value $h/2\pi$ ($k = 1$), while the level 4*d* has the total quantum number 4 and electronic angular momentum of value $3h/2\pi$ ($k = 3$).

It will be observed that the maximum value of *k* is given by the relation $k \leq n$.

The figures in the first column on the left hand side of figure 1 give the value of the spectral terms in wave-numbers ($\nu = 1/\lambda$ in cm^{-1}). Hence the wave-number of any line in the spectrum is given by the difference between the wave-numbers of the initial and final terms (or "levels"). Values of ν are shown for three of the lines, with the corresponding values of λ in Ångströms.¹

¹ $1 \text{ Å} = 10^{-8} \text{ cm}$. Thus $\lambda = 10^8/\nu$. The corresponding frequency of radiation is given by the relation, $\nu^* = 3 \times 10^{15}\nu$. The symbol ν is used in the following discussion to designate wave-number.

The second column in the diagram gives the corresponding value of the energy levels in volts (V), where

$$V = \frac{\nu}{8100}$$

Thus the ionizing potential for Na, $V_i = 5.13$ volts, and the first resonance potential, $V_r = 5.13 - 3.01 = 2.12$ volts.

RYDBERG-RITZ FORMULAS FOR SPECTRAL SERIES (3) (4) (5) (6)

For the spectral series of the alkali metals (and also of the alkaline earth metals) it was observed by Rydberg, Ritz and others that the terms in the series could be represented by expressions which will be recognized as modifications of equation (4). Thus the expression derived by Rydberg is of the form

$$\nu_m = \frac{R}{(m + \alpha)^2} \quad (6)$$

where R denotes the Rydberg constant,² and α is a constant (or approximately so) for all members of the same series, while m assumes successive integral values for the different members of the same series. The values of m assigned in accordance with this relation by spectroscopists are for most series (as will be shown subsequently) less than the values of the total quantum numbers which it has been found necessary to assign to these terms in accordance with Bohr's theory of electronic configurations in atoms. Since the values of n and k are the significant ones from the standpoint of Bohr's theory, it has also become customary to replace the above equation by the expression

$$= \frac{R}{(n - \alpha_K)^2} \quad (7a)$$

$$= R/(\nu_{\text{eff}})^2 \quad (7b)$$

² As mentioned already the value of this constant varies slightly for different atoms. The exact relation is given by $R = R_\infty (1 - m_e/M)$ where m_e/M = ratio of mass of electron to that of the atom, and $R_\infty = 109,737.1$. For He, $R = 109,722.1$ and for H, $R = 109,677.7$.

where n is the total quantum number, and α_K is a constant for each term which depends upon the value of k for the series. The expression $n - \alpha_K$ has also been designated as the "effective" quantum number, n_{eff} , while α_K is known as the "quantum defect."

If now we calculate the values of n_{eff} for the lowest terms in the four spectral series of the alkali metals and compare these

TABLE 1

ATOM		s	p	d	f
H	n	1.00	2.00	3.00	4.00
Li	n	2.00	2.00	3.00	4.00
	n_{eff}	1.59	1.96	3.00	4.00
	α_K	0.41	0.04	0.00	0.00
Na	n	3.00	3.00	3.00	4.00
	n_{eff}	1.63	2.12	2.99	4.00
	α_K	1.37	0.88	0.01	0.00
K	n	4.00	4.00	3.00	4.00
	n_{eff}	1.77	2.23	2.85	3.99
	α_K	2.23	1.77	0.15	0.01
Rb	n	5.00	5.00	3.00	4.00
	n_{eff}	1.80	2.28	2.77	3.99
	α_K	3.20	2.72	0.23	0.01
Cs	n	6.00	6.00	3.00	4.00
	n_{eff}	1.87	2.33	2.55	3.98
	α_K	4.13	3.67	0.45	0.02

values with those of n as derived from considerations of electronic distribution (see subsequent section) we obtain the results shown in table 1 (4) (7). In accordance with (7b) the values of n_{eff} have been obtained from the wave-numbers of the terms by the relation

$$n_{\text{eff}} = \sqrt{109,722/\nu}$$

It will be observed that the differences between the two values of n are less for p than for s and become still smaller for d and f

terms. This observation led Bohr to the interpretation that s orbits penetrate much closer to the nucleus than p orbits, and the latter penetrate nearer than d or f orbits. In fact, in the case of the latter it was considered that the orbit of the valence electron is so far outside the orbits of the other electrons in the atom that the whole system may be regarded as "hydrogen-like." This is evident from the fact that n_{eff} for lowest d and f terms in all the alkali spectra have practically the same values as those observed in the case of H .

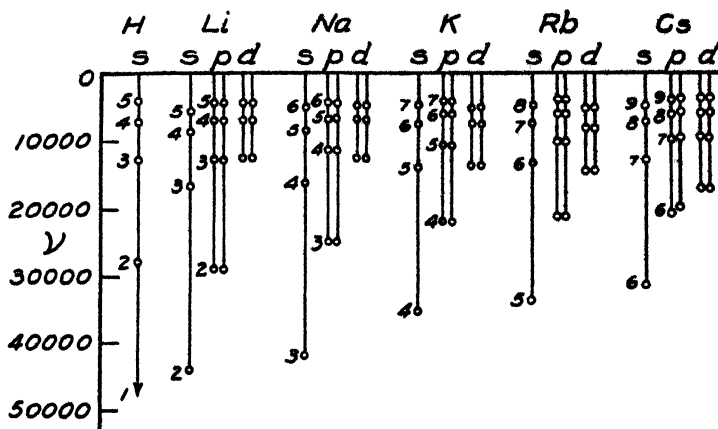


FIG. 2. ENERGY LEVELS FOR ALKALIES COMPARED WITH THOSE FOR HYDROGEN

Figure 2, taken from Hund's book (4) shows graphically the various terms in the spectra of the alkali metals compared with those for H . The numbers attached to each term give the corresponding value of n , while the value of n_{eff} may be obtained from that value of n for H which has the same value of ν .³

MULTIPLY LEVELS

A closer examination of the spectrum of sodium and the other alkali metals shows, however, that the p , d , f levels are really of a

³ The diagram shows the doublet p levels which are mentioned in the next section.

doublet nature. The level indicated as $3p$ in the diagram (figure 1) consists actually of two levels, designated as $3p_1$ and $3p_2$ respectively, with the respective wave-numbers 24492.8 and 24475.65. The transitions between these levels and the normal ($3S$) level give rise to the two lines of the well-known sodium doublet $\lambda 5889.96$ and $\lambda 5895.93$.

In the case of calcium and other metals of group II, it has been found that there are two separate systems of s , p , d , f levels. In one system, the p , d , f levels are triplets, while in the other they are singlets.

As we proceed in the periodic table, from group I to group VIII, it has been observed that the spectral terms exhibit periodic variations such as shown in table 2 for the elements from K ($N = 19$) to Ni ($N = 28$).

It will be observed that firstly, the multiplicity alternates between even and odd, and, secondly, the complexity increases to a maximum in the case of Mn and then decreases. It is the presence of these multiplet levels that gives rise to the extremely complex nature of a large number of spectra, such as those of Ti, Cr, Mn, Fe and other elements. For instance, H. N. Russell has analyzed the spectrum of titanium and finds that it consists of seventeen series divided between singlet, triplet and quintet systems.

Formally the existence of these multiplets may be correlated with the assignment of a *third*, or so-called *inner quantum* number j , which differs for each component of the multiplet term, and which has been interpreted physically as designating the number of units of angular momentum (in terms of $h/2\pi$) of the atom as a whole about an axis of symmetry. According to the older theory it was furthermore assumed that this angular momentum is to be regarded as a resultant of the angular momentum of the orbit and that of the rest of the atom (or kernel).

This led to difficulties in the interpretation not only of the effect of magnetic fields on spectral lines, but also of multiplet terms with more than two components.

The hypothesis of the spinning electron introduced by G. E. Uhlenbeck and S. Goudsmit (11) has, however, served to remove

TABLE 2

K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni
Doublets	Singlets	Doublets Quartets	Singlets	Doublets	Triplets	Quartets Sextets Octets	Triplets	Doublets	Singlets
	Triplets		Triplets	Quartets	Quintets		Quintets	Quartets	Triplets
			Quintets	Sextets	Septets		Septets	Sextets	Quintets

TABLE 3

	SINGLET	DOUBLET	TRIPLET	QUARTET
s	0	1/2	1	3/2
p	1	1/2, 3/2	0, 1, 2	1/2, 3/2, 5/2, 7/2
d	2	3/2, 5/2	1, 2, 3	1/2, 3/2, 5/2, 7/2
f	3	5/2, 7/2	2, 3, 4	3/2, 5/2, 7/2, 9/2
g	4	7/2, 9/2	4, 5, 6	5/2, 7/2, 9/2, 11/2
	QUINTET	SEXTET	SEPTET	OCTET
s	2	5/2	3	7/2
p	1, 2, 3	3/2, 5/2, 7/2	2, 3, 4	5/2, 7/2, 9/2
d	0, 1, 2, 3, 4	1/2, 3/2, 5/2, 7/2, 9/2	1, 2, 3, 4, 5	3/2, 5/2, 7/2, 9/2, 11/2
f	1, 2, 3, 4, 5	1/2, 3/2, 5/2, 7/2, 9/2, 11/2	0, 1, 2, 3, 4, 5, 6	1/2, 3/2, 5/2, 7/2, 9/2, 11/2, 13/2
g	2, 3, 4, 5, 6	3/2, 5/2, 7/2, 9/2, 11/2, 13/2	1, 2, 3, 4, 5, 6, 7	1/2, 3/2, 5/2, 7/2, 9/2, 11/2, 13/2, 15/2

most of these difficulties. According to this theory, each electron in the atom possesses an angular momentum of magnitude $1/2$ ($h/2\pi$). It is, therefore, assigned a quantum number, $s = 1/2$, giving the number of units of angular momentum (in terms of $h/2\pi$).

In the case of the alkali atoms where only one electron is present in the "valence shell," the vector representing the angular momentum of the atom as a whole, j , is the resultant of the vector, s , for the electron, and a vector $l = k - 1$ for the electron in its orbit. Since $s = 1/2$, j can have only the two values $l + 1/2$, and $l - 1/2$. This conclusion is in accordance with the observations on the spectroscopic nature of the normal term for the atoms of the rare gases.

The fact that s has to be assumed as equal to $1/2$ and not to an integral value is necessitated by the existence of doublets in the alkali spectra. For the behavior of the rare gas atoms leads to the conclusion that j for these atoms must be zero. Now if we assume that an electron has $s = 1$, then the resultant values of j for the alkali atom would be $k + 1$, k , and $k - 1$ (i.e., 2, 1, and 0 for the s -state). But the alkali atoms have *no* triplet levels. In fact triplet levels occur in atoms which have an *even* number of valence electrons.

Furthermore, in the s -state of the alkali atoms we have only *one* level. Therefore, the angular momentum of the electron in the orbit cannot be 1, since this would lead to the values $1/2$ and $3/2$ for the resultant value of j (assuming $s = 1/2$). Thus the observed multiplicities are best correlated by assuming that $s = 1/2$, and $l = 0$ for the s -state, $l = 1$ for the p -state, etc., where l designates the number of units of angular momentum for the *electronic orbit*, and takes the place of the quantum number k introduced previously.

On this basis, the j -values assigned to the different component levels in the spectra of the alkali metals are those given in table 3 for a doublet system.

As in the case of the quantum number, k , (or l in the new system) a *selection principle* also governs transitions between different multiplet levels. Not only must we have Δk (or Δl)

$= \pm 1$, but also $\Delta j = \pm 1$ or 0, with the added condition that no transitions can occur from $j = 0$ to $j = 0$.

Denoting the components by the corresponding j -values as subscripts, we find the following transitions:

$p_{1/2} \rightarrow s_{1/2}$; $p_{3/2} \rightarrow s_{1/2}$	two lines
$d_{5/2} \rightarrow p_{3/2}$; $d_{3/2} \rightarrow p_{3/2}$; $d_{3/2} \rightarrow p_{1/2}$	three lines
$f_{7/2} \rightarrow d_{5/2}$; $f_{5/2} \rightarrow d_{5/2}$; $f_{5/2} \rightarrow d_{3/2}$	three lines

The D -lines in sodium correspond to the two $p \rightarrow s$ transitions. Since the energy levels corresponding to $d_{5/2}$ and $d_{3/2}$ are much closer than those corresponding to $p_{3/2}$ and $p_{1/2}$, the two lines ending on the $p_{3/2}$ level are usually merged into one, so that the resulting spectral lines are ordinarily seen as two lines very close together.

For singlet systems and higher degrees of multiplicity than two the values of j for the various component terms are those given in table 3. While the reasons for assigning these values of j to the different terms will be discussed in a subsequent section, it should be noted in the present connection that the following empirical relations may be deduced.

1. The s -level is always single. Denoting the *degree of multiplicity* by r ($r = 1$ for singlets, $r = 2$ for doublets, etc.) and the value of j for this level by j_s .

$$j_s = \frac{r-1}{2}$$

2. For other levels, the possible values of j are given by the series

$$j_s + l, j_s + l - 1, \dots, |j_s - l|$$

where $|j_s - l|$ indicates the absolute difference between the two values. In other words, the values of j are the resultants of l and j_s treated as vectors.

SPECTRAL TERMS AND ELECTRONIC LEVELS

It is very important in this connection to emphasize the difference between spectral terms and electronic levels. In the case

of the alkali metals where we are dealing with transitions of a *single electron* between different orbits, the quantum numbers n, k, j for the orbits are the *same as those of the spectral terms*. But as will be shown in a subsequent section, when there are two or more electrons present in the valence shell (as in case of *Be, B, C*, etc.) the spectral term is the resultant of interactions between the different electrons, each of which is in a different orbit or level. Thus the interactions of two electrons similar to the valence electron in the alkali metals give rise to singlet and triplet terms in the spectra of the alkaline earth metals. In a similar manner the higher degrees of multiplets arise from the interaction of several electrons.

The spectral notation adopted in accordance with these ideas is that first suggested by Russell and Saunders (8) which is as follows:

The *series* is denoted by the capital letters *S, P, D, F, H, I*; the *system* by an index at the upper left hand, giving the *multiplicity*, and the *component* of a multiple term by a subscript on the right, giving the inner quantum number. In the case of terms due to a single electron or where the several electrons all have the same total quantum number, the latter is denoted by an integer prefixed to the main symbol.

For example, in the Na spectrum the 3 p_2 component of the doublet level is designated on the new basis as $3^2P_{3/2}$. However, in order to avoid writing fractional subscripts it is more usual to write the symbol as 3^2P_2 , thus writing $j + \frac{1}{2}$ instead of j in the case of doublets, quartets, sextets, octets, etc. Similarly, in the case of Hg, the lowest level is a singlet ($^1S_0, \nu = 84178.5$). The next higher level is a triplet, and the three component levels are designated as $^3P_2, \nu = 40138.3$, $^3P_1, \nu = 44768.9$, and $^3P_0, \nu = 46536.2$. The well known line in the ultra-violet, $\lambda, 2536.52$, originates in a transition from the 3P_1 level to the 1S_0 and is therefore designated as $^1S_0 - ^3P_1$. It will be observed that the integer corresponding to total quantum number has been omitted in the symbols for these levels. As pointed out by Russell and Saunders, the particular value of n used depends in a number of cases (as will be shown subsequently) on the particular arrangement of

electronic orbits assumed for the model of the atom, and may therefore be modified as a result of further investigation. On the other hand, the behavior of the line λ 2536.5 in magnetic fields (Zeeman effect) and other observations leave no doubt that this line represents a transition from the triplet to a singlet level as indicated above. Furthermore, as is evident from an inspection of the energy diagram for Na, the total quantum number, n , has no bearing on the possibility or impossibility of a transition between two given levels.

NORMAL AND ANOMALOUS ZEEMAN EFFECTS (4) (5) (9) (10)

In 1896 Zeeman discovered that the lines of series spectra are influenced by magnetic means. In the simplest case there appear instead of one line, when viewed longitudinally, that is, when the ray is in the direction of the magnetic lines of force, two lines (Zeeman doublet, longitudinal effect), but when viewed transversely, that is, when the ray is perpendicular to the magnetic lines of force, instead of one line, three lines are observed (Zeeman triplet, transverse effect). Of the latter three lines one occupies the position of the original unresolved line, and the other two are displaced by equal amounts to greater or smaller wave-lengths, and occupy the same position in the spectrum as the two lines of the doublet in the longitudinal effect. The displacement amounts to

$$\Delta \nu_n = \frac{e}{2m_0c} \cdot \frac{H}{2\pi} \quad (8)$$

$$= 1.41 \times 10^6 H$$

where $\Delta \nu_n$ = change in frequency

e/m_0 = ratio of charge to mass of electron in e. s. u.

H = intensity of field in gauss (37).

The resolution into three components with separation given by the above equation is known as the *normal* Zeeman effect. The magnetic field in this case leads to a change in the energy level by an amount

$$\pm h \cdot \Delta \nu_n = \pm \frac{eH}{2m_0c} \cdot \frac{h}{2\pi} = \pm o \quad (9)$$

where ω is known as the "Larmor precession," since the change in frequency is interpreted physically as a frequency of precession of the orbit about the direction of the magnetic lines of force.

Most spectral lines, however, exhibit a much more complex resolution in magnetic field. Instead of three components in a transverse magnetic field there are observed a larger number, and the separations from the original position may in all cases be represented by rational multiples (not necessarily integral) of the normal value, $\Delta\nu_n$. This is known as the *anomalous Zeeman effect*.

In accordance with Bohr's point of view we must assume that the magnetic field affects both the initial and final levels involved in the origin of any given line, but the effect on each level depends upon the values of the two quantum numbers, j and k of the particular orbit (or level). In presence of a magnetic field, an atomic system having the resultant angular momentum $j \cdot h/2\pi$ can orient itself with respect to the magnetic field only in certain definite positions (space quantization in a magnetic field), and to each of these positions there corresponds a definite magnetic moment of the atomic system.

MAGNETIC QUANTUM NUMBER, m

The development of this point of view leads to the conclusion that for a level of inner quantum number, j , in presence of a magnetic field there exists a magnetic moment, m , and that the resultant change in the value of the energy level is given by $mg\lambda_n$, where $m = j, j-1, \dots, -(j-1), -j$.

The constant, g is known as the "splitting-up" factor, since it gives the *distance* in multiples of $\Delta\nu_n$ between the different components into which the energy level is resolved in the magnetic field, while the number of components is obviously equal to the total number of different values which m may assume. The negative values are due to the fact that the magnetic moment of the orbit may orient itself either with the field or in the opposite direction.

The value of g is determined by the values of k and j of the level. From these considerations there readily follow the various

generalizations regarding the Zeeman effect which were observed previously to the development of the theory as sketched above. The most important of these is that the Zeeman effect is similar for lines which correspond to transitions between levels of the same initial quantum numbers, k_i and j_i , and the same final quantum numbers, k_f and j_f . Thus lines belonging to the same spectral series show similar values of m and g . Consequently it is possible *from an investigation of the behavior of spectral lines in magnetic fields, to determine the values of k and j (i.e. the series type and multiplet component) for the corresponding energy levels.* This conclusion has been of greatest importance in revealing the nature of the terms in the case of the more complex spectra, such as these of Mn, Cr and Ru.

TABLE 4

SPECTRAL LEVEL	k	l	j	CORRESPONDING VALUES OF m	TOTAL NUMBER OF LEVELS IN MAGNETIC FIELD
S	1	0	$\frac{1}{2}$	$\pm \frac{1}{2}$	2
P	2	1	$\frac{1}{2}; \frac{3}{2}$	$\pm \frac{1}{2}; \pm \frac{1}{2}; \pm \frac{3}{2}$	6
D	3	2	$\frac{3}{2}; \frac{5}{2}$	$\pm \frac{1}{2}; \pm \frac{3}{2}; \pm \frac{5}{2}$	10
			$\frac{5}{2}; \frac{7}{2}$	$\pm \frac{1}{2}; \pm \frac{3}{2}; \pm \frac{5}{2}; \pm \frac{7}{2}$	
F	4	3	$\frac{5}{2}; \frac{7}{2}$	$\pm \frac{1}{2}; \pm \frac{3}{2}; \pm \frac{5}{2}; \pm \frac{7}{2}$	14
			$\frac{7}{2}; \frac{9}{2}$	$\pm \frac{1}{2}; \pm \frac{3}{2}; \pm \frac{5}{2}; \pm \frac{7}{2}; \pm \frac{9}{2}$	

The existence of the anomalous Zeeman effect has been most satisfactorily interpreted physically by the hypothesis of the spinning electron, which has been mentioned in a previous section, together with an additional hypothesis, also introduced by Uhlenbeck and Goudsmit, that the magnetic moment corresponding to the spin vector, $s = 1/2$, is *not* $1/2 \left(\frac{e}{2m_0c} \right) \left(\frac{h}{2\pi} \right)$ as would follow from Larmor's theory, but *double* this quantity, that is equal to $2s$ in units $\left(\frac{e}{2m_0c} \right) \left(\frac{h}{2\pi} \right)$. It is the presence of this factor 2 that leads to the "anomalous" effect.

Since the magnetic moment of the electron may be in the same

direction as, or opposed to, the direction of the external magnetic field, it is obvious that the term 2S_1 will yield two terms in presence of a magnetic field. In the case of the doublet P terms, one term, 2P_1 , has $j = 1/2$, while the other, 2P_2 has $j = 3/2$. Hence in presence of a magnetic field, the values of m due to $j = 1/2$ are $m = 1/2, -1/2$, while those due to $j = 3/2$ are $m = 3/2, 1/2, -1/2, -3/2$. Thus the two P -levels give rise to six levels in a magnetic field. Similar considerations lead to the values given in table 4 for the total number of terms or levels which are due to different values of l in presence of a magnetic field.

Thus the total number of possible states of an atomic system for given value of l , in a weak magnetic field, is equal to $2(2l+1)$.

PASCHEN-BACK EFFECT (4) (9) (10)

In weak magnetic fields the number of orientations is determined by first obtaining the resultant momentum, j , (in units $h/2\pi$) of the momenta l and s and then assigning to the vector j all the possible integral values ranging from j to $-j$.

This procedure is logical only when the magnetic field is so weak that it does not affect the constraining forces between the valence or "optical" electron (that is, the electron which is passing from one orbit to another) and the rest of the atom ("Rumpf"). But in stronger magnetic fields these constraining forces are loosened, and finally, in very strong fields, the vectors l and s will no longer influence each other. The stationary states of the atomic system will then correspond to states of the model in which l and s are oriented separately, so that the component m_s of s is either parallel or antiparallel with the field, while the component m_l of l will orient itself so that $m_l = l, l-1, \dots, -(l-1), -l$. Taking into account that the magnetic moment due to s , the angular momentum of spin of the electron, is equal to $2s$, while that due to l is m_l (both moments being measured in units $\left(\frac{e}{2m_0c} \cdot \frac{h}{2\pi}\right)$), the magnetic quantum number in very strong is given by $m' = m_l + 2m_s$.

This is the explanation of the phenomenon observed by Paschen and Back (9) that in strong magnetic fields both the number and spacing of the components of a given term are different from those obtained in weak magnetic fields.

STONER'S AND MAIN SMITH'S SCHEME OF ELECTRON DISTRIBUTION

As shown in table 4 the total number of orbits in a magnetic field for an electron of serial quantum number l is given by $2(2l+1)$. Now corresponding to any given value of n , values of l may range from $n-1$ to 0. Hence the number of orbits (or levels) for any

TABLE 5

n	VALUE OF l	SPECTRAL DESIGNATION	NUMBER OF LEVELS IN MAGNETIC FIELD
1	0	1s	2
2	0	2s	2
	1	2p	6
3	0	3s	2
	1	3p	6
	2	3d	10
4	0	4s	2
	1	4p	6
	2	4d	10
	3	4f	14

given value of n will also be limited in accordance with the following scheme (table 5) which may be extended for larger values of n .

Thus the maximum number of levels for a 1s, 2s, 3s or ns electron are 2, while for an nf electron (where n denotes the total quantum number and f the particular series type) the maximum number of levels is 14.

Now according to E. C. Stoner (12) and Main-Smith (13) the numbers given in the last column of table 5 must also represent the *maximum number of electrons of that spectral type which an atomic system can possess*. As will be shown in a subsequent section a real basis for this hypothesis is furnished by Pauli's

principle. However, it is possible, even without discussing this principle, to follow up the logical consequences of the hypothesis in its application to the problem of electron distribution in atoms.

The electron in the atom of H is in orbit of type $1s$. To form the atom of He we add a second electron (and, of course, increase the charge on the nucleus by one unit of positive electricity). Since two electrons of type $1s$ may exist in an atom, the second electron is also of this type. Table 6 shows the type of orbit formed by each electron as it is added in succession to build up the atoms of higher atomic number.

With the addition of a third electron we enter upon the second period (Li-Ne), and the first 2 electrons belong to the $2s$ group. Since this represents the maximum number for a $2s$ group the third electron will be of type $2p$, and this last group will be filled with the eighth electron (Ne = 10).

The eleventh electron, corresponding to Na ($N = 11$) and the twelfth (Mg, $N = 12$) form the $3s$ group, while Al ($N = 13$) completes the $3p$ group, and thus resembles Ne in the configuration of the external group of eight electrons.

When we come to K ($N = 19$) the analogy with Na and Li leads to the conclusion that the added electron must enter a $4s$ group. Much more important, however, in the case of this and the other atoms is a criterion which points definitely to the same conclusion and which may be stated as follows: During the process of formation of the atoms by addition of successive electrons the *last bound electron always assumes that orbit for which the energy of the atomic system is a minimum*. Thus in the case of K the spectrum shows that an electron in a $3d$ orbit has a higher energy value than in a $4s$. The same holds true for Ca, but in the case of Sc ($N = 21$), the energy in a $3d$ is lower than that in a $4p$ level (since the $4s$ level is already filled). Consequently we begin, with this element, to fill up the $3d$ group, which can contain 10 electrons as a maximum. As shown in table 6 this group is completed in Cu, and the $4p$ group is completed in Kr.

Proceeding to still larger values of N we find that there is a constant tendency for electrons to shift between levels such as $4d$ and $5s$, $4f$, $5d$ and $6s$, etc. To understand the reasons for

TABLE 6

ELEMENT	N	K	L		M			N		NORMAL TERM	V _i	
		1s	2s	2p	3s	3p	3d	4s	4p			
H	1	1								² S ₁	13.54	
He	2	2								¹ S ₀	24.48	
Li	3	2	1							² S ₁	5 37	
Be	4	2	2							¹ S ₀	9.48	
B	5	2	2	1						² P ₁	8.4	
C	6	2	2	2						³ P ₀	11 24	
N	7	2	2	3						⁴ S ₂	12.2	
O	8	2	2	4						³ P ₂	13 56	
F	9	2	2	5						³ P ₂	18 6	
Ne	10	2	2	6						S ₀	21 5	
Na	11	Neon "Shell"			1					² S ₁	5 12	
Mg	12				2						¹ S ₀	7 61
Al	13				2	1					² P ₁	5 96
Si	14				2	2					³ P ₀	10 19
P	15				2	3					⁴ S ₂	13 3
S	16				2	4					³ P ₂	10 20
Cl	17				2	5					³ P ₂	13 0
A	18				2	6					¹ S ₀	15 48
K	19	Argon "Shell"						1		² S ₁	4 32	
Ca	20							2		¹ S ₀	6 09	
Sc	21							1	2	² D ₂	6 57	
Ti	22							2	2	³ P ₂	6 80	
V	23							3	2	⁴ F ₂	6 76	
Cr	24							5	1	⁷ S ₃	6 74	
Mn	25							5	2	⁶ S ₃	7 40	
Fe	26							6	2	⁵ D ₄	7 83	
CO	27							7	2	⁴ F ₃	7 81	
Ni	28							8	2	³ F ₄	7 64	
Cu	29	Argon "Shell"				10	1		² S ₁	7 69		
Zn	30					10	2		¹ S ₀	9 35		
Ga	31					10	2	1	² P ₁	5 97		
Ge	32					10	2	2	³ P ₀			
As	33					10	2	3	⁴ S ₂	11.5		
Se	34					10	2	4	³ P ₂	12.7		
Br	35					10	2	5	² P ₂			
Kr	36					10	2	6	¹ S ₁	13 3		

TABLE 6—Continued

ELEMENT	N	NUMBER IN INNER SHELLS	N		O			P	NORMAL TERM	V _i
			4d	4f	5s	5p	5d	6s		
Rb	37	Kryp- ton Shell			1				² S ₁	4 16
Sr	38				2				¹ S ₀	5 67
Y	39		1		2				² D ₂	
Zr	40		2		2				³ F ₂	
Cb	41		4		1				⁶ D ₁	
Mo	42		5		1				¹ S ₃	7 35
Ma	43		6		1				⁶ D ₆	
Ru	44		7		1				⁵ F ₅	7 7
Rh	45		8		1				⁴ F ₆	7 7
Pd	46		10						¹ S ₀	8 5
Ag	47	Palladium "Shell"			1				² S ₁	7 54
Cd	48				2				¹ S ₀	8 95
In	49				2	1			² P ₁	5 76
Sn	50				2	2			³ F ₀	7 37
Sb	51				2	3			⁴ S ₂	8 5
Te	52				2	4			³ P ₂	
I	53				2	5			² P ₂	
Xe	54				2	6			¹ S ₀	11 5
Cs	55	(1s-4d = 46)			(5s-5p = 8)			1	² S ₁	3 88
Ba	56							2	¹ S ₀	5 19
La	57				(5s-5p = 8)		1	2	² D ₂	
Ce	58		1				1	2	³ H ₄	
Pr	59		2				1	2	⁴ K ₆	
Nd	60		3				1	2	⁵ L ₆	
Il	61		4				1	2	⁶ L ₅	
Sa	62		5				1	2	⁷ K ₄	
Eu	63		6				1	2	⁸ H ₂	
Gd	64		7				1	2	⁹ D ₂	
Tb	65		8				1	2	⁸ H ₉	
Ds	66		9				1	2	⁷ K ₁₀	
Ho	67		10				1	2	⁶ L ₁₀	
Er	68		11				1	2	⁵ L ₁₀	
Tu	69		12				1	2	⁴ K ₉	
Yb	70		13				1	2	³ H ₆	
Lu	71		14				1	2	² D ₂	

TABLE 6—*Concluded*

ELEMENT	N	NUMBER IN INNER SHELLS	O		P			Q	NORMAL TERM	V ₁
			5d	5f	6s	6p	6d	7s		
Hf	72	(1s-5p = 68)	2		2				² F ₃	
Ta	73		3		2				⁴ F ₁	
W	74		4		2				¹ D ₀	
Re	75		5		2				⁶ S ₅	
			6		1				⁶ D ₄	
Os	76		6		2				⁵ D ₄	
			7		1				⁵ F ₅	
Ir	77		7		2				⁴ F ₅	
			8		1				⁴ F ₅	
Pt	78		8		2				³ F ₄	
			9		1				³ D ₃	
			10						¹ S ₀	
Au	79	(1s-5d = 78)			1				² S ₁	9 20
Hg	80				2				¹ S ₀	10 39
Tl	81				2	1			² P ₁	6 08
Pb	82				2	2			² P ₀	7 39
Bi	83				2	3			⁴ S ₂	8.0
Po	84				2	4			³ P ₂	
	85				2	5			² P ₂	
Rn	86				2	6			¹ S ₀	
Ra	87	(1s-5d = 78)			(6s-6p = 8)			1	² S ₁	
	88							2	¹ S ₀	
Ac	89						1	2	² D ₂	
Th	90			1			1	2	³ H ₄	
							2	2	³ F ₂	
Ux	91			2			1	2	⁴ K ₄	
							3	2	⁴ F ₂	
Ur	92			3			1	2	⁶ L ₆	
							4	2	⁶ D ₀	

adopting the particular distribution assigned to each element it has been necessary to investigate not only the ordinary or optical spectra, but also the x-ray spectra. We must therefore digress in this connection to a discussion of the important deductions which have been derived from investigations of spectra of this class.

X-RAY SPECTRA. *K, L, M, N, O LEVELS* (5) (14)

Since x-ray spectra may be considered as indicating the process by which an atom undergoes reorganization after a disturbance in its interior, it is evident that a study of the corresponding energy levels must lead to information about the distribution of electrons in the interior of atoms. A detailed study of x-ray spectra from the point of view of the Bohr theory has thus led to the identification of a number of different sets of levels which are listed in table 7, together with the corresponding values of n , l , and j . The last line gives the analogous levels in the optical spectra of the alkali metals.

TABLE 7
Classification of x-ray levels

LEVEL	<i>K</i>	<i>L</i>			<i>M</i>				
SUB-LEVEL...		<i>L_I</i>	<i>L_{II}</i>	<i>L_{III}</i>	<i>M_I</i>	<i>M_{II}</i>	<i>M_{III}</i>	<i>M_{IV}</i>	<i>M_V</i>
<i>n</i>	1	2	2	2	3	3	3	3	3
<i>l</i>	0	0	1	1	0	1	1	2	2
<i>j</i>	1	1/2	1/2	3/2	1/2	1/2	3/2	3/2	5/2
Optical analog	1 <i>S</i>	2 <i>S</i>	2 <i>P₁</i>	2 <i>P₂</i>	3 <i>S</i>	3 <i>P₁</i>	3 <i>P₂</i>	3 <i>D₂</i>	3 <i>D₃</i>

In the elements of higher atomic number, there are also present *seven N* levels ($n = 4$, $l = 0, 1, 2, 3$) and *five O* levels ($n = 5$, $l = 0, 1, 2$).

Figure 3 taken from the paper by Bohr and Coster (14) shows the values of $\sqrt{\nu/R}$ (ν = wave-number; R = Rydberg constant) for the different levels plotted against the corresponding values of N . It will be observed that for any given type of level the value of $\sqrt{\nu/R}$ varies approximately linearly with N (Moseley's law), so that $\sqrt{\frac{\nu}{R}} = \frac{N-\sigma}{n}$, where σ is a "screening" constant, and n the total quantum number corresponding to the level. Applying this relation to the curves shown in Fig. 3, it is found that as a first approximation $n = 1$ for the *K* levels, $n = 2$ for the *L* levels and $n = 3$ for the *M* levels. By analogy with the optical levels in hydrogen it has therefore been concluded

that the values of n for K , L and M levels must be 1, 2 and 3 respectively. In fact, it has been shown by Millikan that the

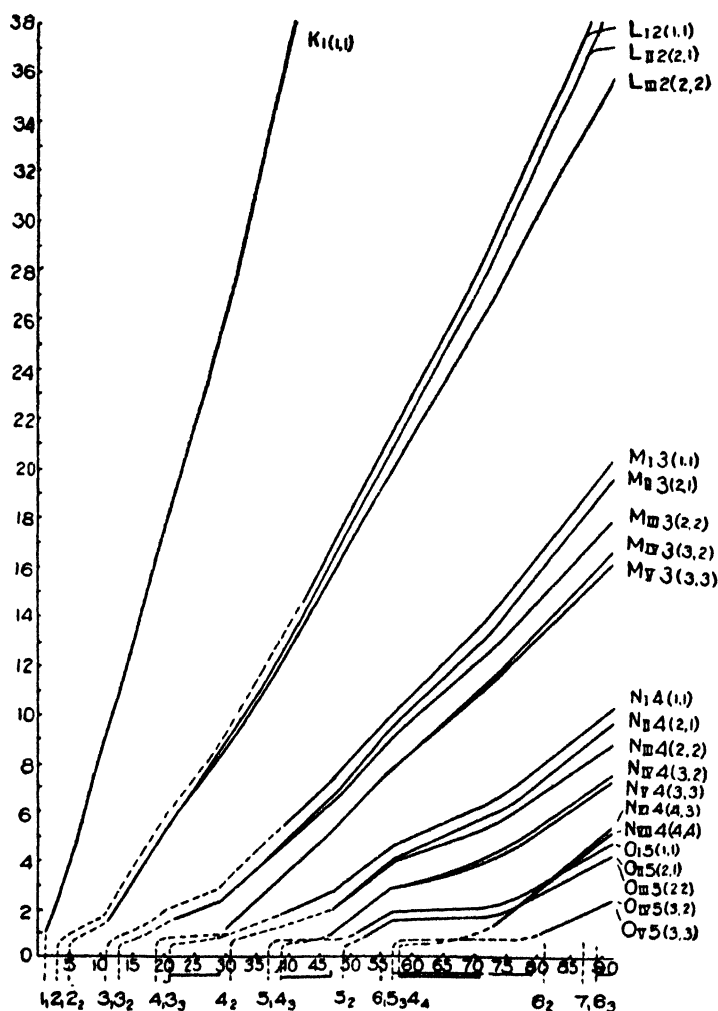


FIG. 3. MOSELEY DIAGRAM FOR X-RAY LEVELS

limiting wave-number in the Lyman series for H (corresponding to the normal level in H) forms an extension of the K level.

The regular increase in $\sqrt{\nu/R}$ with N for any one series of levels, so noticeable in the diagram, and first observed by Moseley, is obviously different from that observed for say S levels in optical spectra. The reason for this difference is due to the difference in origin of x-ray lines and optical lines. While the latter are due to transitions from *virtual* orbits external to the outermost normal orbits of the electrons in the atom, x-ray lines indicate a re-arrangement of electrons in the interior of the atom. Thus the K_α , K_β and K_γ lines are due to a transition of an electron from L , M and N levels respectively to the K level.

The fact that the values of the x-ray energy levels correspond to total quantum numbers—1, 2, 3 for K , L and M levels respectively lends strong support to Bohr's conception of the mode of formation of atoms by addition of electrons in successive levels of higher energy value. Furthermore, a closer study of the values of N at which certain x-ray levels begin to appear, furnishes evidence regarding the stages at which a re-arrangement occurs in the external electrons between such levels as $3d$ and $4s$ or $4f$ and $5d$.

This conception of Bohr's regarding the mode of formation of atoms may now be described in more detail. In the H atom the electron is normally in an orbit for which $n = 1, l = 0, j = 1/2$, but in excited states the electron may be raised to any one of the other (*virtual*) levels corresponding to higher values of n and the accompanying possible values of l and j ($l = n - 1, n - 2, \dots, 1, 0; j = l \pm 1/2$) as shown in table 5 for the values $n = 1, 2, 3$, and 4. The optical spectrum of H results from transitions between these levels as well as the normal level. As more electrons are added these enter in succession the higher (*virtual*) levels for H , and the number of electrons which can enter each of these levels is given by Stoner's rule. Owing to the increased nuclear charge the energy for these levels is, of course, greater than that for the same level in the H atom. X-ray spectra are produced, as stated already, by reorganization of the electrons in completed levels, while optical spectra are produced by removal of an electron from one of the external levels to a still higher virtual level.

Figure 4, also taken from the paper by Bohr and Coster, shows the energy levels for the rare gas radon ($N = 86$), as deduced from

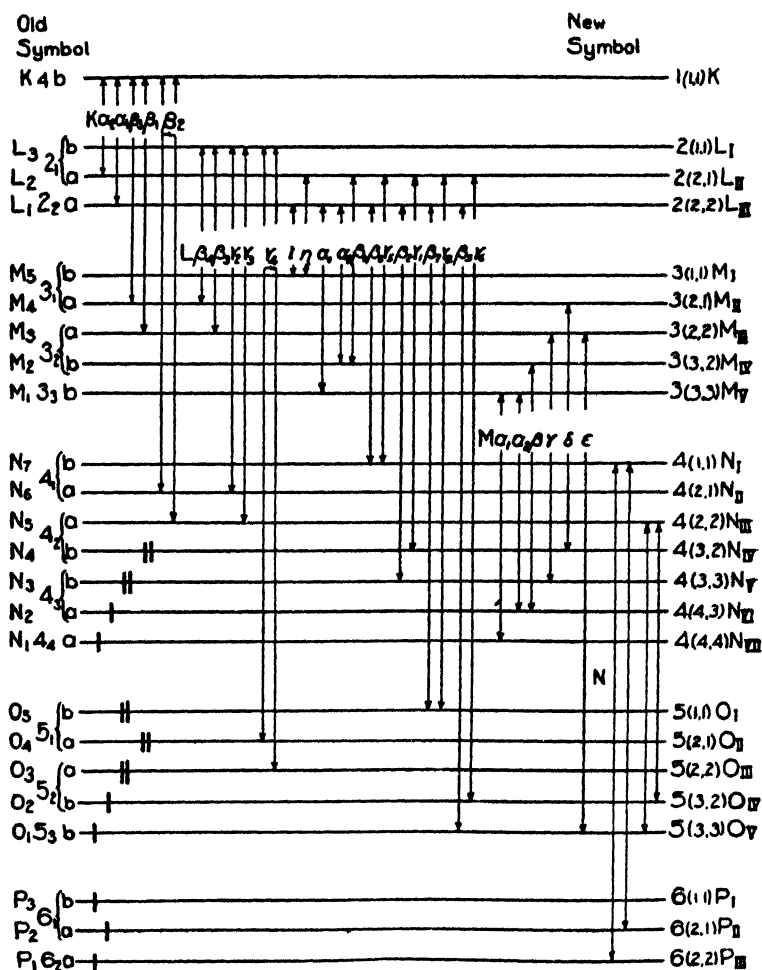


FIG. 4. X-RAY LEVELS IN RADON

x-ray absorption data for elements of high atomic number. The symbols suggested by Bohr and Coster are given on the right hand side together with the values of n , k , and $j + 1/2$, while the

older nomenclature is given on the left. The arrows indicate possible transitions and correspond therefore to observed lines in the x-ray spectra. For elements of lower atomic number the number of levels obviously decreases. For instance, in the case of Cs ($N = 55$) the L and M levels are present and N_I to N_V as well as O_I , O_{II} and O_{III} . The valence electron occupies the $6s$ level. In the case of K, the K and L levels are filled, and the M_I , M_{II} and M_{III} ($3s$, $3p_1$, and $3p_2$ respectively) while the valence electron occupies the N_I or $4s$ level. For the latter atom all the other levels indicated in the diagram are *virtual* levels to which the valence electron may be raised, thus leading to the formation of lines in the optical region when the electron returns to the $4s$ level.

DISTRIBUTION OF ELECTRONS IN ATOMS

We can now consider in further detail the electron distribution for the different atoms as shown in table 6. It is evident that the periodic recurrence of atoms with similar physical and chemical properties is correlated with similar energy levels for the external electrons. The gases of Group O exhibit the arrangement of eight outer electrons of which two are in s levels and six in p levels. A convenient notation for this, also adopted in recent papers, is s^2p^6 , the power index indicating the number of electrons in the particular level. Similarly we can write for the electrons external to the rare gas "shell" the following symbols for different groups:

Group I	Li, Na, K, Rb, Cs	s
Group IIa	Be, Mg, Ca, Sr, Ba, Ra	s^2
Group IIb	Zn, Cd, Hg	$d^{10}s^2$
Group IIIa	B, Al, Ga, In, Tl	s^2p
Group IIIb	Sc, Y, La, Ac	ds^2
Group VII	F, Cl, Br, I	s^2p^5

The rather anomalous properties of the Ni, Pd and Pt groups, as well as the properties of the elements immediately following them (Cu, Ag and Au) are to be interpreted by the fact that in the first-named elements a d level either becomes completed or tends to be completed, (and in a great many respects the elements

immediately following a completed level behave like those which follow the elements of Group O). We shall however discuss this point again in connection with magnetic and spectroscopic properties of the atoms.

The position and number of the 14 rare earth elements Ce to Lu is of special interest. Firstly, this arrangement shows that there cannot exist more or less than 14 such similar elements. Secondly the fact that the added elements from Ce to Lu enter into a 4 *f* level while the 5 *d* and 6 *s* levels remain unaltered, is in agreement with the well-known fact that all these 14 elements are strikingly similar in all their chemical and physical properties. It is indeed remarkable that both La, which precedes Ce, and Hf which follows Lu are so much more definite in their characteristic properties that there can be no question as to which elements they each resemble. Thus La falls naturally in the Sc, Y group, while Hf belongs just as naturally with Ti and Zr.

There is some doubt regarding the exact distribution of electrons in the elements beyond Hf and in table 6 there have been incorporated the various arrangements suggested for each atom by J. S. McLennan, A. B. McLay and H. G. Smith (15). Further investigation is necessary before a definite decision on these questions can be reached. Thus, depending upon whether Pt resembles Ni or Pd, the distribution in 5*d* and 6*s* levels of the last three electrons will vary from d^8s^2 (Ni) to d^{10} (Pd).

Again, in the case of Th, and the next two elements, it is doubtful whether the added electrons enter the 5 *d* level (in analogy with Ti and Zr etc.) or 5 *f* level (in analogy with Ce, etc.).

DISTRIBUTION OF ELECTRONS IN SUB-LEVELS

One more point remains to be mentioned in connection with the arrangement of electrons in atoms shown in table 6. According to the evidence obtained from x-ray spectra (see Table 7 and figure 3) the levels which we have designated as *p*, *d*, *f*, etc. in table 6 are each to be regarded as doublet levels, corresponding to two different values of *j*, the inner quantum number. Table 7 shows the analogy (discussed more fully in a subsequent section) between these sub-levels derived from x-ray data with the optical sub-levels in the spectra of the alkali metals.

Now, as has been pointed out in connection with table 4, the number of electrons in each completed level is equal to $2(2l+1) = 2\Sigma(j + \frac{1}{2})$, that is, to the sum of the two values of $j + \frac{1}{2}$ for each of the levels p , d , and f . Stoner (12) has accordingly suggested that the number of electrons associated with each sub-level separately is also equal to $2(j + \frac{1}{2})$. In support of this assumption he points out that this number represents the observed term multiplicity in a weak magnetic field. That is, as shown in table 4, the number of different values of m is equivalent to $2(j + \frac{1}{2})$ and therefore these states of the atomic system are "always

TABLE 8

LEVEL	SPECTRAL NOTATION	SUB-LEVEL			CORRESPONDING X-RAY DESIGNATION	TOTAL NUMBER OF ELECTRONS IN SUB-LEVEL
		n_{k_j}	l	j		
$s \dots \dots \dots$	$^2S_{\frac{1}{2}}$	n_{11}	0	$\frac{1}{2}$	$K_1 L_1 M_1 N_1 O_1 P_1$	2
$p \dots \dots \dots$	$^2P_{\frac{1}{2}}$	n_{21}	1	$\frac{1}{2}$	$L_{II} M_{II} N_{II} O_{II} P_{II}$	2) 4)
	$^2P_{3/2}$	n_{22}	1	$\frac{3}{2}$	$L_{III} M_{III} N_{III} O_{III}$	
$d \dots \dots$	$^2D_{3/2}$	n_{32}	2	$\frac{3}{2}$	$M_{IV} N_{IV} O_{IV}$	4) 6)
	$^2D_{5/2}$	n_{33}	2	$\frac{5}{2}$	$M_V N_V O_V$	
$f \dots \dots$	$^2F_{5/2}$	n_{43}	3	$\frac{5}{2}$	N_{VI}	6) 8)
	$^2F_{7/2}$	n_{44}	3	$\frac{7}{2}$	N_{VII}	

possible and equally probable, but only manifest themselves separately in the presence of the external field. Electrons can enter a group until all possible orbits are occupied, when the atom will possess a symmetrical structure."

According to this view therefore the maximum number of electrons in each of the sub-levels would be as shown in table 8.

Only those levels have been mentioned in this table, which are actually found completed in the periodic arrangement. While the numbers given in the last column of table 8 undoubtedly represent the distribution of electrons in the atoms of Group O, there is still considerable discussion regarding the exact distribu-

tion between each sub-level for a large number of the elements. This is specially true of the elements in Groups V, VI and VII and VIII as well as the rare earth group. Under these circumstances it has been considered advisable to defer discussion of this topic to a subsequent section of the paper.

SIMULTANEOUS TRANSITIONS OF TWO OR MORE ELECTRONS

In a previous section the origin of doublet levels in the spectra of the alkali metals was accounted for by assuming an interaction

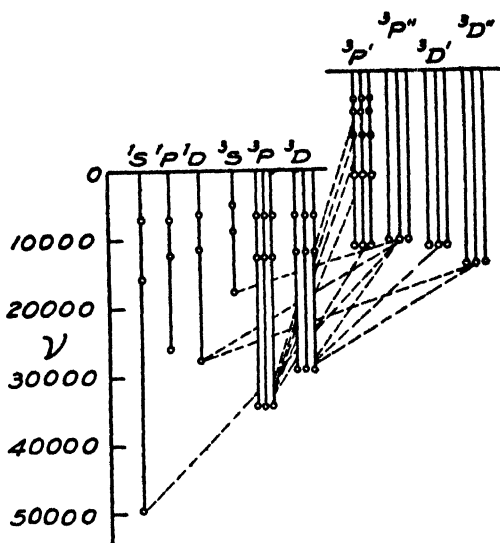


FIG. 5. NORMAL AND NEGATIVE TERMS FOR CALCIUM

between the angular momentum of the spin of the electron ($s = \frac{1}{2}$) and the angular momentum l of the electronic orbit. In this manner the values of j , the resultant momentum, given in table 4, were obtained. Evidently the assumption that the spectral terms are due to changes in the nature of the orbit of a single electron is not adequate to account for higher multiplet levels than the doublet. In other words, transitions of a single valence or "optical" electron between different levels does not explain the mode of formation of more complex spectra. Although

Heisenberg, Landé and others had formerly derived certain rules by which they could calculate the nature of the multiplet levels formed by successive addition of electrons, there existed, as mentioned previously, certain theoretical difficulties in the whole scheme. While the conception of the spinning electron has served to remove some of these difficulties, the evidence brought forward by Russell and Saunders (8) for the simultaneous transition of two electrons has helped materially in clarifying the whole problem of complex spectra.

It had been known for some time that in the spectrum of Ca there exist not only the two systems of singlet and triplet levels mentioned already, but also an anomalous group of levels for which the limiting terms (corresponding to the ionizing potential in the normal spectra of alkali and alkaline earth metals) have *negative* values. Figure 5 (4) which gives the energy diagram for Ca illustrates what is meant by negative terms. The S , 3S and other normal terms converge to the term marked 0 which corresponds to the ionizing potential. On the other hand there exist a group of terms such as $^3P'$, $^3P''$, etc., which converge to a higher term than that corresponding to the ionizing potential, so that some of the terms are negative (with reference to the convergence limit for the normal terms.)

The clue to the nature of this phenomenon is afforded by the observation that the difference between the limit of the normal term and that of the anomalous terms is equal to the difference between the lowest and next highest terms in the spectrum of ionized Ca (Ca^+ or CaII). The latter spectrum resembles that of K , since the atom has the same total number of electrons, external to the nucleus, in both cases. Thus the anomalous terms must obviously be due to *simultaneous transitions* of one of the two valence electrons from the $4s$ to a $3d$ level and of the other electron from the $4s$ to any other level. When *both electrons* return to lower levels a *monochromatic radiation* is emitted just as in the case of the transitions of a single electron.

Formally the observed phenomena may be accounted for on the following basis, which forms an extension of the vector-model used in interpreting the relations between l and s for the terms in the spectra of the alkalis.

The l -vectors of the two electrons interact to form a resultant L -vector, similarly the s -vectors combine to form an S -vector, and the resultant of L and S gives rise to a j -vector for the particular term. As an illustration of the application of this method of procedure, let us consider the case where $l_1 = 0$, and $l_2 = 2$ (as in that of Ca, where one of the electrons is in a $4s$ and the other may be in a $3d$ level). The resultant level has $L = 2$ (D term) while $S = 1$ or 0 , corresponding to a singlet or triplet level. Thus we obtain $^1D'$ and $^3D'$ terms. Similarly when $l_1 = 1$ and $l_2 = 2$, it may be shown that the resultant terms are $^3F'$, $^3D'$, $^1D'$ and $^1S'$. The dashes on the upper right hand side of the letters indicate that these terms are "anomalous."

DERIVATION OF TERMS IN COMPLEX SPECTRA

On the basis of these considerations it is now possible to discuss briefly the method developed by F. Hund (4) (16), for deriving the types of spectral terms which may result from the interaction of more than one electron in an atom. Before proceeding with this it is well to summarize briefly the discussion in the previous sections regarding quantum numbers. Following Russell's notation we distinguish the *state of a single electron* in an atom from the *state of the atomic system which corresponds to a definite spectral term*.

1. The state of a *single* electron is ordinarily designated by *four* quantum numbers: (a) n , the principal or total quantum number, which determines to a first approximation the energy of the state; (b) l , the "serial" quantum number, since it serves to designate the system to which the state belongs (s , p , d , f , etc.); (c) m_l , the magnetic quantum number which designates the orientation of the orbit in a magnetic field; (d) m_s , the magnetic quantum number which designates the direction of spin in a magnetic field.

It is to be observed, furthermore, that

$$\begin{aligned} l &= n - 1, n - 2, \dots 0 \\ m_l &= l, l - 1, \dots -(l - 1), -l \\ m_s &= \pm 1/2 \end{aligned}$$

and

$$m = m_l + m_s$$

in weak magnetic fields.

2. The *state of an atomic system* is defined by three numbers: (a) the "serial" quantum number, L , which is the resultant of the l -vectors for the individual electrons, and which designates the type of system (S, P, D, F , etc.); (b) the spin quantum number S which corresponds to the resultant of the spin vectors of the individual electrons, and (c) the inner quantum number, j , which represents the total number of units of angular momentum of the atomic system for the given state, and which is the resultant (vectorially) of L and S .

The value of S determines the multiplicity, r , of the system which is given by the relation

$$r = 2S + 1.$$

Hence $r = 1$ (singlets) for $S = 0$; $r = 2$ (doublets) for $S = 1/2$, and so forth.

As mentioned previously, the multiplicity is indicated by an index at the upper left hand corner of the symbol for the series; while the value of j is shown by an index at the lower right hand corner.

According to the method used by Hund, the values of L , S and j for the different possible terms resulting from the interaction of two or more electrons are determined as follows: Forming the sums of the different combinations of values of m_l for the individual electrons, we obtain values of $\Sigma m_l = m_L$; similarly values of $\Sigma m_s = m_s$ are obtained for each different combination. From these the values of $m = m_L + m_s$ are derived and these, in turn, are used to deduce the possible values of j , since corresponding to a given value of the latter in a spectral term, the magnetic quantum numbers, m , are given by the series of values $j, j-1, \dots -j$. Similarly corresponding to a given value of L in the spectral terms, the values of m_L are given by $L, L-1, \dots -L$, and corresponding to given value of S , the values of m_s are given by $S, S-1, \dots -S$.

Hund furthermore assumes that of all the possible spectral terms thus obtained, the *normal term*, or term corresponding to the lowest energy state of the atomic system is determined in accordance with the following two rules:

(1) The normal term corresponds to the highest value of the multiplicity, r .

(2) For regular multiplets the lowest value, $j = L - S$, and for inverted multiplets the highest value, $j = L + S$, is chosen.

From table 4 it follows that in case of levels which contain their maximum number of electrons, $\Sigma m = 0$. Consequently

TABLE 9
 $n_2 = n_1 = 3; l_1 = l_2 = 2$

m_{l_1}	m_{l_2}	m_{s_1}	m_{s_2}	m_L	m_S	m
2	2	$+\frac{1}{2}$	$-\frac{1}{2}$	4	0	4
	1	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	3	100-1	4332
	0	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	2	100-1	3221
	-1	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	1	100-1	2110
	-2	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	0	100-1	100-1
1	1	$+\frac{1}{2}$	$-\frac{1}{2}$	2	0	2
	0	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	1	100-1	2110
	-1	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	0	100-1	100-1
	-2	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	-1	100-1	0-1-1-2
0	0	$+\frac{1}{2}$	$-\frac{1}{2}$	0	0	0
0	-1	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	-1	100-1	0-1-1-2
0	-2	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	-2	100-1	-1-2-2-3
-1	-1	$+\frac{1}{2}$	$-\frac{1}{2}$	-2	0	-2
-1	-2	$\pm\frac{1}{2}$	$\pm\frac{1}{2}$	-3	100-1	-2-3-3-4
-2	-2	$+\frac{1}{2}$	$-\frac{1}{2}$	-4	0	-4

$j = 0$, and the interaction of all these electrons in the *completed level* must lead to a term having the designation 1S_0 . In table 6 the second last column gives the *lowest* or *normal term* in the spectrum of the element, and it will be observed that the 1S_0 term is characteristic of all those elements in which a level becomes completely filled, such as Ne; Mg, Pd and analogous elements.

The method of determining the spectral terms which may result from the interaction of two or more electrons may be illustrated by considering a case, such as that of Ti, in which two ($3d$) elec-

trons are added to the two (4s) electrons already present in the shell external to that of argon. Since the two (4s) electrons represent a completed level, $j = m = 0$ for these and therefore they contribute no angular momentum to the resultant angular momentum of the two (3d) electrons. Designating one of the latter electrons by subscript 1 and the other by subscript 2, we obtain the values of n, l, m_l, m_s given in table 9.

The values of m correspond to $j = 4$ (two terms), $j = 3$ (one term), $j = 2$ (three terms), $j = 1$ (one term) $j = 0$ (two terms). Altogether there result nine terms. The values of $m_s = 0$ indicate singlet terms, while $m_s = 1, 0, -1$ correspond to triplet terms. Also since the maximum value of m_L is 4, we may expect terms in G, F, D, P and S levels. Distributing the nine terms among these levels we obtain the spectral terms $^1S_0, ^1D_2, ^1G_4; ^3P_2, ^3P_1, ^3P_0, ^3F_4, ^3F_3, ^3F_2$. In accordance with the rules given by Hund, the normal term must be of type 3F_2 , the lowest value of j being used since the number of electrons is less than half that corresponding to a completed (3d) level.

It will be observed that for the cases $m_{l1} = m_{l2}$ only the values $m_s = \frac{1}{2}$ and $m_s = -\frac{1}{2}$ have been used. The reason for excluding other combinations of m_s is based on a rule which is known as "*Pauli's Principle of Restriction of Quantum Numbers*," and which may be enunciated as follows:

In any one atom there cannot exist two electrons having the four quantum numbers, (n, l, m_s and m_l) respectively the same in both.

That is, if for two electrons $n_1 = n_2; l_1 = l_2; m_{l1} = m_{l2}$, then m_{s1} cannot be equal to m_{s2} . Furthermore it is obvious that in the case illustrated above, $m_{l1} = 2, m_{l2} = 1$ is identical with $m_{l1} = 1$ and $m_{l2} = 2$, and so forth.

It will be recognized that Pauli's principle is essentially the real justification for Stoner's theory of electron distribution in levels; since, for instance, putting more than 6 electrons in a p level would be in contradiction with this principle.

On the basis of these considerations it is possible to derive, as has been shown by Hund, the different spectral terms which may be obtained from the interaction of two or more electrons having the same values of n and l respectively. The results of these

calculations are given in table 10 for the case $l = 1$ (p - electrons).

Since $\Sigma m_l = \Sigma m_s = \Sigma m = 0$ for all the electrons in a completed level, it follows that 4 equivalent electrons give the same spectral terms as 2, and 5 the same as 1. The only difference occurs in the order of the multiplet levels. While in the case of the first three electrons, the lowest multiplet level corresponds to the least value of j , the order is reversed in the case of the last three, as is shown in the last column of table 10 which gives the multiplet level of lowest energy content and therefore corresponds to the "normal" term.

It will be observed that in table 10 the multiplicity $r = 2$ for one electron changes to $r + 1$ and $r - 1$ on addition of a second

TABLE 10
Terms for equivalent p-electrons

NUMBER OF ELECTRONS	TERMS	LOWEST MULTIPLET
1	2P	$^2P_{1/2}$
2	1S_0 1D_2 3P	3P_0
3	2P 2D 4S	$^4S_{3/2}$
4	1S_0 1D_2 3P	3P_2
5	2P	$^2P_{3/2}$
6	1S_0	1S_0

electron, and the same result is reproduced on addition of a third electron. Then the multiplicities decrease in reverse order as the last three electrons are added. In Table 1 a similar set of observations were shown for the elements of the first long period. A given multiplicity r in any one atom becomes altered to $r + 1$ and $r - 1$ in the succeeding atom in the periodic arrangement. (Heisenberg's "branching" rule.) This leads to increasing complexity in the degree of multiplicity until the number of electrons added corresponds to half of the maximum number possible for the given level, and then the complexity decreases in reverse order. Thus in the $3d$ level the complexity reaches a maximum with Cr. In all cases the maximum occurs in that case for which the number of added p or d electrons is one half the maximum possible number for the completed shell.

On the basis of Hund's theory this phenomenon is accounted for readily since for every added electron, the value of S , the resultant spin vector is decreased or increased by $1/2$, and consequently the multiplicity must be decreased or increased by 1. Furthermore, the occurrence of maximum complexity in the middle of the series corresponding to a given level is due to the same cause as the repetition of spectral terms in the second half of the series which has been mentioned above, i.e., Pauli's exclusion principle.

PERIODIC VARIATIONS IN NATURE OF NORMAL TERM

Table 6 gives the normal term for each of the 92 elements (15) (16). In certain cases (such as Cr, Cu, Ni and Pd) the nature of the normal terms as derived from the Zeeman effect has led to the conclusion that the electron distribution must be that given in the table. In other cases, such as the rare earth group and a large number of the elements of higher atomic number, detailed investigation of the spectra still remains to be carried out before a final decision can be reached regarding the exact electron distribution. However, Hund's theory of the relation between spectral terms and electron arrangement as outlined above, has been so singularly confirmed in such a large number of cases that no doubt exists in regard to its validity.

The periodic variation in spectral terms with increasing atomic number is evident from table 6. Atoms which have a similar arrangement of electrons in the "valence" or outer levels, have identical spectral terms. In accordance with the rules enunciated in a previous section, it is observed that the term corresponding to $m - \chi$ electrons in a level is the same as that for χ electrons, where m denotes the *maximum number* of electrons for the completed level; the only difference which occurs is that in the second half of the group of m atoms, the component of the multiplet level which has the lowest energy value is that corresponding to the highest value of j , ("reversed multiplets") as contrasted with "regular multiplets" at the beginning of the group.

Thus Ti with $(3d)^2$ electrons and Ni with $(3d)^8$ electrons exhibit 3F terms; V with $(3d)^3$ and Co with $(3d)^7$ have 4F terms, and simi-

larly for other elements in the group, Sc to Ni. The only exception is Cu for which the expected electron distribution $(3d)^9 (4s)^2$ does not apparently possess as low an energy content as that corresponding to $(3d)^{10} (4s)$. Hence the latter yields a normal term which is different from that of Sc.

The rare earth group is an illustration of the case for which the completed $4f$ level should contain 14 electrons, and it will be observed that the normal term for $14 - \chi$ electrons is of the same nature (except as regards the particular value of j) as that for χ electrons ($\chi = 1$ to 6). The spectral terms given in table 6 are derived on the assumption that the electron configuration corresponds to $(4f)^n(5d)(6s)^2$ where n is the number of electrons in $4f$ level. On the other hand, Hund has suggested the alternative possibility $(4f)^n(5d)^2(6s)$ in which case the normal terms would be of still higher degree of complexity. Thus for Gd ($N = 64$) the normal term would be of type ${}^{11}F$ instead of 9D as given in the table, and for all the other atoms of the group the normal level would be altered to that corresponding to the next higher value of L with an increase in degree of multiplicity by 2.

In the case of Os, Ir, Pt and some of the other elements more than one distribution is possible and the normal terms have been given for each of these, as calculated by McLennan, McLay and Smith (15). In a recent paper, A. C. Hausmann (18) has, however, reached the conclusion, from an examination of the Zeeman effect on the terms in the arc spectrum of Pt, that the normal term is of type 3D_3 corresponding to the configuration d^9s .

ENHANCED SPECTRA

We shall now consider types of spectra which are of considerable importance inasmuch as they furnish experimental evidence for Bohr's general conception of the building up of atoms by addition of successive electrons. It has already been stated that according to Bohr's theory the energy of an atomic system consisting of a nucleus of charge Ne and a single electron is given by

$$E_n = - \frac{RchN^2}{n^2} \quad (5)$$

where E_n denotes the energy for the state corresponding to total quantum number n .

As a consequence of this relation it was shown by Bohr that certain lines previously ascribed to hydrogen really belong to ionized helium, since in this case $N = 2$. It is evident that for an ionized atom, the number of electrons is one less ($N - 1$) than the charge on the nucleus, and, therefore, for orbits of large quantum number the force between the "optical" electron and the rest of the atom will approximate that of a nucleus for which $N = 2$, acting on a single electron. Thus the spectra of singly ionized atoms will possess terms in which the Rydberg constant, R , obtained in ordinary spectra, is to be replaced by $4R$. Now such spectra had been observed and were known as "*spark spectra*" in contradistinction to the more usual *arc spectra*."

Bohr's theory, however, led to the conclusion that it should also be possible to observe spectra due to atoms in which two or more electrons have been removed, and for these "*enhanced*" spectra the value of the Rydberg constant would have to be replaced by $4R$, $9R$ or p^2R according as one, two or $(p - 1)$ electrons have been removed from the neutral atom. It has become conventional to designate such a "*stripped*" atom by the symbol of the element to which is attached at the lower right hand corner a Roman numeral whose value is equal to that of p . Thus we speak of the spectrum of F_{VII} and signify by this symbol the spectrum emitted by a fluorine atom which has been stripped of *six* external electrons and therefore actually has the same number of electrons as the Li atom.

As an important corollary from the remarks made in connection with Hund's theory of the formation of spectral terms it follows that the first spark spectrum of any element ought to exhibit the same multiplet structures as the arc spectrum of the element preceding it in the periodic table. (Kossell-Sommerfeld's displacement law.) Furthermore, the enhanced spectrum of an element of atomic number N from which p electrons have been removed should resemble the ordinary spectrum of an element of atomic number $N - p$. (But, of course, the coefficient of the Rydberg constant in the case of the enhanced terms would be $(p + 1)^2$.

A very interesting illustration of these considerations is furnished by the spectra of the series Na_I , Mg_{II} , Al_{III} , Si_{IV} , which have been investigated by Paschen, Foote and Fowler (19), and P_V , S_{VI} , and Cl_{VII} investigated by Millikan and Bowen (20). All these spectra exhibit doublet levels, but owing to the effect of in-

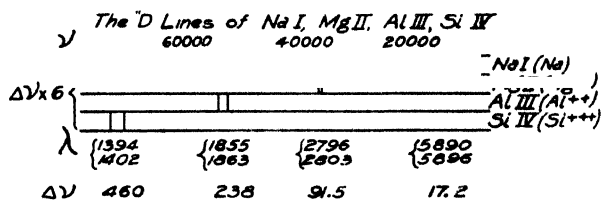


FIG. 6. "D" LINES IN ENHANCED SPECTRA OF SODIUM-LIKE ATOMS

creasing net nuclear charge, the wave lengths of similar lines (i.e., corresponding to transitions between levels of same spectral designation) become shorter as the number of electrons removed from the atom increase. Thus figure 6 (20), (21) shows the "D" lines in the spectra of the first four atoms.

TABLE 11
Comparison of frequencies of series terms for stripped atoms Na_I to S_{VI}

	3S	3P	3D	4P
H	12192 8	12192 8	12192 8	6858 4
Na/1	41449 0	24475 7	12276.2	6860.4
Mg/4	30316 9	21376 6	12444 3	6866 8
Al/9	25494 9	19504 0	12611 0	6871.3
Si/16	22756 8	18272 6	12733 7	6874.2
P/25	20979 7	17401 9	12811 8	6876.4
S/36	19729 6	16753 6	12856 6	6878 1

Table 11 (20) gives a comparison of the lowest series terms for these spectra. In accordance with equation (5) the observed values of ν have been divided by p^2 where $p-1$ designates the number of electrons removed from the neutral atom. The values of ν for the corresponding levels in H are given at the top. Here again it is observed that the values of D and F levels are approximately the same as those of similar levels in H , (thus indicating

that the corresponding orbits are completely external to the rest of the atom) while owing to the deeper penetration of the s and p orbits the corresponding energy values (as measured by the values of ν) are greater. Consequently the quantum defect ($\alpha_K = n - n_{eff}$) decreases as we pass from Na_I to S_{VI} .

Millikan and Bowen have also obtained the spectra of the series L_{II} to F_{VII} and these show variation in the frequencies of series terms similar to those observed for the series Na_I to Cl_{VI} .

While these spectra all resemble those of the alkali metal group in exhibiting doublet levels, Millikan and Bowen have also found spectra of stripped atoms which resemble the metals of Group II in exhibiting singlet and triplet levels. Such are the series:

$$\begin{array}{l} Be_I, B_{II}, C_{III}, N_{IV}, O_V, F_{VI} \\ Mg_I, Al_{II}, Si_{III}, P_{IV}, S_V, Cl_{VI} \end{array}$$

In fact the spectra of all these atoms have been obtained in all the various states of enhancement. Thus, series, such as P_I to P_V ; Si_I to Si_{IV} ; N_I to N_V , have been observed spectroscopically and in each case the expectations based on Hund's theory have been signally confirmed (22). Similar results have been obtained by Russell for the enhanced spectra of Ti (23).

MULTIPLY LEVELS IN STRIPPED ATOMS

In a previous section it was stated that an investigation of x-ray spectra has led to the identification of a number of different sets of levels, which are listed in table 7. Millikan and Bowen (24) pointed out in 1924 that the relations observed for the separations between certain pairs of these levels could be applied to analogous levels in optical spectra. These observations have since been extended to multiplet levels in stripped atoms and furnish additional evidence for the periodic variation in the structure of spectral terms with variation in atomic number.

Now in the case of the L_{II} L_{III} levels it had been shown by A. Sommerfeld (5) that the separation

$$\Delta\nu = 0.365 (N - \sigma)^4 \quad (10)$$

where N = atomic number, and σ denotes the "screening" constant. Similarly for the M_{II} M_{III} levels:

$$\Delta\nu = 0.108 (N - \sigma)^4 \quad (11)$$

These levels constitute what are known as "relativity" or regular doublets and as may be seen from an inspection of figure 3 the separation, $\Delta\nu$, increases rapidly with increase in N . On the other hand, the levels, L_I , L_{II} ; M_I M_{II} from the so-called

TABLE 12

	$2P_1 - 2P_2$			$2S - 2P_1$	
	$\Delta\nu$	$\sqrt{\Delta\nu/305}$	σ	$\Delta\nu$	$\sqrt{\nu_S} - \sqrt{\nu_{P_1}}$
Li_I	0 34	0 981	2 019	14903 8	39 470
Be_{II}	6 61	2 063	1 937	31927 6	44 297
B_{III}	34 4	3 116	1 884	48358 7	45 594
C_{IV}	107.4	4 142	1 858	64481 2	46 068
N_V	259 1	5.162	1 939	80455 1	46 276
O_{VI}	533 8	6 184	1 816	96367.8	
	$3P_1 - 3P_2$			$3S - 3P_1$	
	$\Delta\nu$	$\sqrt{\Delta\nu/108}$	σ	$\Delta\nu$	Diff.
Na_I	17 18	3 550	7 450	16956.17	
Mg_{II}	91 55	5 394	6 606	35669 42	18713 25
Al_{III}	234 00	6 820	6 180	53681 53	18012 11
Si_{IV}	461 84	8 084	5 916	71285 49	17603 96
P_V	794 82	9 259	5 741	88649.42	17363 93
S_{VI}	1267 10	10 404	5 596	105866.04	17216 62
Cl_{VII}	1889 5	11 496	5.504	123001 23	17135 19

"irregular" or "screening" doublets. For these levels the values of the differences, $\sqrt{\nu_{L_I}} - \sqrt{\nu_{L_{II}}}$ and $\sqrt{\nu_{M_I}} - \sqrt{\nu_{M_{II}}}$ increase only slightly with increase in N , as shown in figure 3.

According to Hertz this relation indicates that the difference between the values of σ , the screening constant, for the L_I (or $2s$) and L_{II} (or $2p_1$) levels is practically constant for all the atoms.

As a result of recent developments in the theory of wave mechanics Sommerfeld's physical interpretation of the difference, $\Delta\nu$, between the wave numbers of the two levels L_{II} , L_I , etc. has been discarded (25) and it has been shown that the observed

differences must be regarded as due to differences in direction of spin of the electron for the two levels whose inner quantum numbers are $j = l + 1/2$ and $j = l - 1/2$ respectively. But it is of interest that the relation derived on the basis of this theory is identical in the case of doublet levels with Sommerfeld's equations (10) and (11).

Table 12⁴ gives the values of $\Delta\nu$ and σ for doublet terms in the spectra of stripped atoms of types Li_I and Na_I .

Thus the levels $2P_2$ and $2P_1$ are analogous to L_{II} and L_{III} respectively; $3P_2$ and $3P_1$ to M_{II} and M_{III} respectively, while $2S$

TABLE 13
Analogous levels

X-RAY LEVEL	OPTICAL LEVEL	X-RAY LEVEL	OPTICAL LEVEL
K	$1S$	$\{N_I$	$4S$
$\{L_I$	$2S$	$N_{II}\}$	$4P_1$
$\{L_{II}$	$2P_1$	$\{N_{III}\}$	$4P_2$
L_{III}	$2P_2$	$\{N_{IV}\}$	$4D_2$
$\{M_I$	$3S$	$N_V\}$	$4D_3$
$\{M_{II}$	$3P_1$	$\{N_{VI}\}$	$4F_3$
$\{M_{III}\}$	$3P_2$	$N_{VII}\}$	$4F_4$
$\{M_{IV}\}$	$3D_2$	$\{O_I$	$5S$
M_V	$3D_3$	$\{O_{II}$	$5P_1$
		$O_{III}\}$	$5P_2$

and $3S$ are analogous to L_I and M_I respectively. The analogy has been found to extend over the whole range of x-ray and optical levels as shown in table 13.

Irregular doublet levels are shown bracketted on the left hand side, while regular doublet levels are bracketted on the right.

The analogy is best illustrated by the Moseley diagram shown in figure 7 taken from the paper by Bowen and Millikan (20).

Their remarks in this connection are worth quoting. They state:

This whole comparison may be very beautifully made graphic in precisely the way in which Moseley first exhibited the relations

⁴ The subscripts to the P levels are those to be assigned on the basis of the j values and are the reverse of those used formerly by spectroscopists.

between atomic numbers and x-ray spectra. Indeed, no more striking demonstration of the fact that x-ray laws hold throughout the *optical* region as well can be made than is found in figure 7. In this figure we have plotted all our stripped atom levels, corresponding to the second row of the periodic table, in a Moseley diagram. That these optical levels follow the *Moseley law* is exhibited by the linear relation everywhere found between $\sqrt{\nu}/R$ and atomic number. That they follow the irregular doublet law is shown by the parallelism of the 3s, 3p and 3d lines, all of which have constant screening differences,

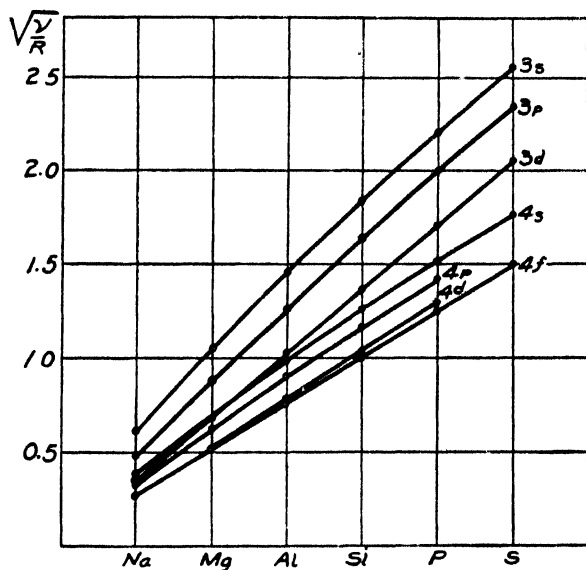


FIG. 7 MOSELEY DIAGRAM FOR OPTICAL LEVELS IN SODIUM-LIKE ATOMS

corresponding as they do to 3_1 , 3_2 and 3_3 orbits,⁵ respectively. Quite similarly, the irregular doublet law appears in the near parallelism of the 4s, 4p, 4d and 4f lines, which correspond to 4_1 , 4_2 , 4_3 , and 4_4 orbits, respectively. . . . The diagram thus furnishes a very satisfactory graphical illustration of the extension of two of the three x-ray laws, namely, the Moseley law and the irregular doublet law, into the field of optics. That it fails to illustrate the extension of the third, or regular doublet law, is merely due to the fact that the regular

⁵ The subscripts here refer to the values of $k = l + 1$ for the orbits.

doublet separations are so small as not to be visible at all upon a drawing of this scale.

As shown in table 12 the value of the screening constant, σ , is approximately 2 for the series Li_I to O_{VI} , which corresponds to the number of electrons in the K shell, and thus represents the residual number of electrons after the L electrons are removed. In the case of the series Na_I to Cl_{VII} , the value of σ is approximately 6, instead of the expected value 10, which is interpreted as due to

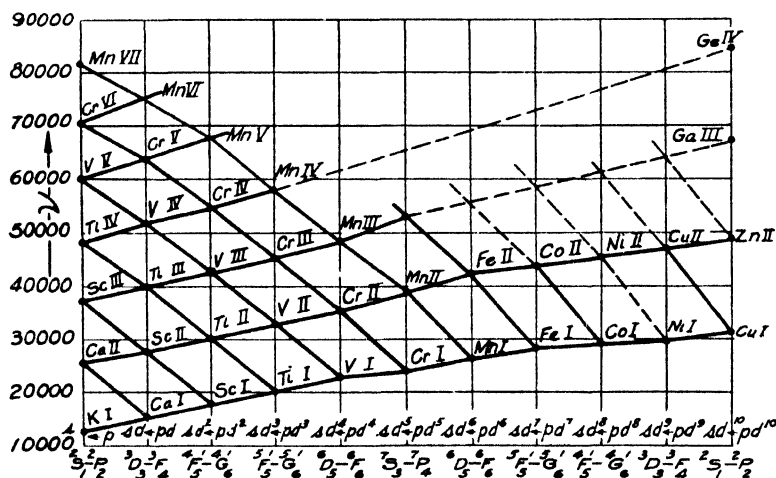


FIG. 8 ANALOGOUS LINES IN ENHANCED SPECTRA OF ELEMENTS OF FIRST LONG PERIOD

a penetration of the $3P$ orbits within the L shell, and, indeed, when σ is calculated from the doublet separations $3D_2 - 3D_3$ and $4F_3 - 4F_4$, the value obtained is approximately 10.

R. C. Gibbs and H. E. White have investigated, in a similar manner to Bowen and Millikan, the relations existing for doublets of stripped atoms of the potassium, rubidium, and caesium types (26). The irregular and regular doublet laws were found to be also applicable to these atoms. Figure 8 taken from a more recent paper (27) shows the regular displacement of multiplets for the elements in the first long period. Vertical lines represent atoms having the same number of external electrons* ("iso-elec-

* A very apt term used, as far as the writer knows, only by Gibbs and White.

tronic systems"). The ordinates give the values of ν for the spectral lines indicated at the bottom of the figure. Thus, along the first vertical column are given the frequencies of the line $4^2S_1 - 4^2P_1$ (first line of the principal series) for the potassium-like atoms K_I to Mn_{VII} , which have only one valence electron. The second vertical line corresponds to the calcium-like atoms, Ca_I to Mn_{VI} , which are two electron systems, and the ordinates give the frequencies of the spectral line whose initial level is the 3F_4 level (corresponding to a combination of one electron in a p level and the other in a d level) and final level is the 3D_3 level (whose origin is one electron in an s level and the other in a p level). Above the spectral designation of each line are given the combinations of electrons which give rise to the corresponding levels.

It will be observed, as has been mentioned previously, that the multiplicity alternates between even and odd values as we pass from K-like to Cu-like atoms, and also that the degree of this multiplicity passes through a maximum (seven) for the atoms for which the number of electrons and their distribution are the same as that of Cr_I . This variation in multiplicity is in agreement with the fact that as we pass from K-like to Cu-like atoms the number of electrons in the atomic system external to the argon shell increases from one at the beginning to eleven at the end, and, as is to be expected on the basis of Hund's theory, atomic systems having the same number and distribution of electrons external to the next lower rare gas exhibit the same spectral structure. In fact, this figure illustrates beautifully all that has been stated in the previous discussion in regard to the conclusions which can be drawn from Hund's theory.

As a further illustration it is of interest to refer briefly to the various spectra of Ti which have been investigated in detail by Russell (23). The systems of terms present in the successive spark spectra have been found by him to be as follows:

	T_I	T_{II}	T_{III}	T_{IV}
Systems present in spectrum.....	Quintets Triplets Singlets	Quartets Doublets .	Triplets Singlets	Doublets
Electron configurations.	$(3d)^2 (4s)^2$	$(3d)^2 (4s)$	$(3d)^2$	$(3d)$

IONIZING POTENTIALS OF NEUTRAL ATOMS

The ionizing potential has already been defined as the energy required to remove one of the external electrons completely from the rest of the atomic system, thus leaving the latter in the state of a singly ionized atom. Hence the ionizing potential may ordinarily be calculated from the wave-number of the lowest energy level by the relation

$$V_i = \frac{\nu}{8100}$$

H. N. Russell (28) has, however, pointed out that in an atomic system in which there are two or more valence electrons (as in Groups II to VIII) there should be more than one ionizing potential corresponding to different processes for the removal of the electron.

As an illustration he considers the elements of the first long period (K to Ni). If we designate the number of electrons outside the "argon shell" by n , the normal state of the atom corresponds to the configuration s in the case of K and $d^{n-2}s^2$ in all the other cases (except Cr for which the normal configuration is d^5s).

Now, as is evident from table 2 the degree of multiplicity, r , changes by $+1$ or -1 as we pass from one element to the next in the periodic arrangement (Heisenberg's "branching" rule.) Therefore, in the removal of one of the $4s$ electrons from the configuration $d^{n-2}s$ the resulting ion may have a greater or less multiplicity than the neutral atom. Corresponding to these two possibilities we have two different ionizing potentials:

- A. $d^{n-2}s^2$ to $d^{n-2}s$ (greater multiplicity)
 B: $d^{n-2}s^2$ to $d^{n-2}s$ (smaller multiplicity)

Thus in the case of Ca the normal level is of type 1S_0 . On ionization there results a potassium-like ion, in which the lowest term is of greater multiplicity, i.e., of type 2S_1 . On the other hand, the ionization may be accompanied by the simultaneous transition of the second electron from the $4s$ to the $3d$ level, and

again the multiplicity may be either increased or decreased, with the corresponding two possible ionizing potentials:

$$\begin{aligned} C: d^n - 2s^2 \text{ to } d^{n-1} & \text{ (greater multiplicity)} \\ D: d^n - 2s^2 \text{ to } d^{n-1} & \text{ (smaller multiplicity)} \end{aligned}$$

Russell therefore defines the *principal ionizing potential* as that which represents the difference in energy between the normal states of the atom and that of the ion. He furthermore accounts for the difference between the four ionizing potentials *A*, *B*, *C* and *D* as follows:

The differences between the ionization potentials *A* and *B*, or *C* and *D* admit of a simple explanation: The reverse process for the transition *C* and *D*, whether in the arc or the spark, consists of adding a 4s electron to a partially completed shell of 3 *d* electrons. In case *C*, the multiplicity is increased, that is, the added electron is *spinning in the same direction* as the resultant spin of those already present. In case *D*, it is spinning in the opposite direction. Now, as Hund has pointed out, the magnetic interaction produces attraction in the first case, and repulsion in the second, so that the energy of binding is greater for *C* than for *D*.

Thus the principal ionizing potentials for Ca, Sc, Ti, Mn, Fe and Zn correspond to a transition of type *A*; while those for K, Cr and Cu correspond to a transition of type *B*. For V, Co and Ni the transition involving the least energy change is of type *C*.

In the last column in table 6 the ionizing potentials have been given for all the elements for which reliable data are available (29). As far as possible the principal ionizing potentials, as defined by Russell, have been given. Figure 9 shows the values plotted as a function of the atomic number, and as in the case of other properties it is observed that the ionizing potential is to a large extent distinctly a periodic function of *N*.

Thus the inert gases occupy the highest places, while the chemically very active alkali metals are located at minimum points on the curve. The extremely electronegative elements F and Cl also have values of V_i which are relatively close to those of group O. In a general way, the ionizing potential tends to increase as the number of electrons in any given level is increased, and reaches

a maximum value when the particular level is completely filled. This is illustrated by the relative values of V , for the triads Li, Be, B; Na, Mg, Al; Zn, Ga, Ge; Ag, Cd, In; and Au, Hg, Tl. In each case the element in the middle corresponds to the completion of an s level and with the last named element electrons begin to enter a p level. On the other hand, it is difficult to interpret the fact that V , for sulfur is less than for either of the neighboring elements.

The similar appearance of the two sections of the graph corresponding to the K-Zn series on the one hand and to the Rb-Cd

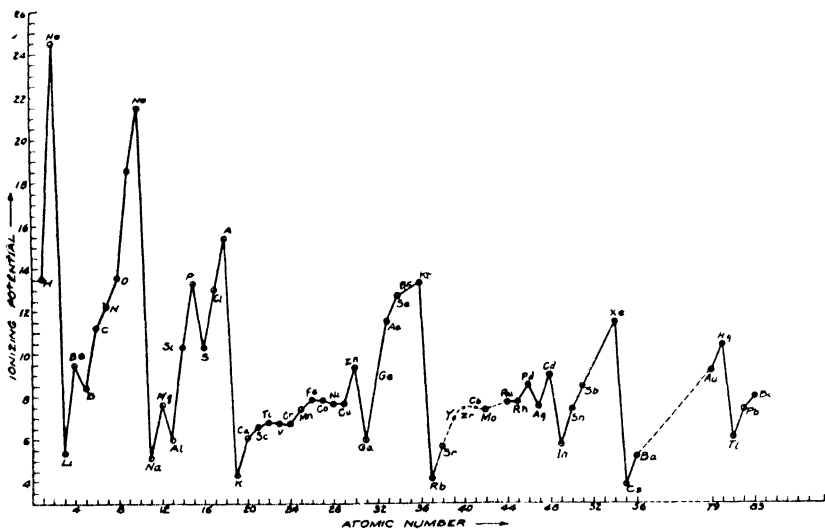


FIG. 9 PLOT OF IONIZATION POTENTIAL AS FUNCTION OF ATOMIC NUMBER

series on the other is quite evident. Also it will be observed that for elements belonging to the same group, V , decreases as N increases. In a general way it is possible to prophesy from these data the ionizing potentials for the other elements of the periodic arrangement. The graph shows the interpolated values for Y, Zr and Cb which have been located by analogy with the corresponding points for Sc, Ti and V, and undoubtedly, La, Ce and the rare earth elements will occupy similar positions with respect to Ba, with values of V , ranging between 7 and 8, reaching a low

TABLE 14
Ionizing energy of stripped atoms⁷

ELEMENT	N	ELECTRON DISTRIBUTION IN VALENCE SHELL	ENERGY OF REMOVAL OF OUTER ELECTRON FORM						
			M	M ⁺	M ⁺⁺	M ⁺⁺	M ⁺⁺	M ⁺⁺	M ⁺⁺
H	1	s	1 00	∞					
He	2	s ²	1 80	4 00	∞				
Li	3	s	0 40	5 60	9 00	∞			
Be	4	s ²	0 70	1 34	11 4	16 0	∞		
B	5	s ² p	0 62	1 79	2 80	19 3	25 0	∞	
C	6	s ² p ²	0 83	1 79	3 36	4 74	20 0		
N	7	s ² p ³	1 07	2 18	3 49	5 43	7 19		
O	8	s ² p ⁴	1 00	2 58	4 05	5 69	8 07	10 14	
F	9	s ² p ⁵	1 25	2 39	4 62	6 40	7.56	(11 0) ⁸	13 60
Ne	10	s ² p ⁶	1 59	2 70					
Na	11	s	0 38	3 50					
Mg	12	s ²	0 56	1 11	5 98				
Al	13	s ² p	0 44	1 34	2 09	9 00			
Si	14	s ² p ²	0 76	1 20	2 34	3 32	12.47		
P	15	s ² p ³	0 98	1 46	2 22	3 53	4 78		
S	16	s ² p ⁴	0.76	1 74	2 37	3 48	4 93	6 47	
Cl	17	s ² p ⁵	0 96	1 66	2 76	3.50	5 00	(6 8) ⁸	8 40
A	18	s ² p ⁶	1 16	2.12 ⁹					
K	19	s	0 32	2 34					
Ca	20	s ²	0 45	0 87	3 76				
Sc	21	s ² d	0 49	0 95	1 82	5 33			
Ti	22	s ² d ²	0 50	1 00	2.04	3.18	7 07		

⁷ The energy is expressed as a multiple of 13 54 V (= 312,368 cal. per gram atom; = 109,678 cm⁻¹).

⁸ Values obtained by interpolation on log-log plot. The values 7 56 for F and 3.50 for Cl are apparently low.

⁹ The value for A⁺ is from a paper by Compton, Russell and Boyce, Proc. Nat. Acad. (in print).

maximum value at Hf and then decreasing to W with another pronounced maximum at about 10 volts in Pt.

IONIZING POTENTIALS OF STRIPPED ATOMS

As a result of the work of Millikan and Bowen, A. Fowler, H. N. Russell and others, our knowledge of the nature and energy values of levels in enhanced spectra has been increased enormously in the past three years. While certain results of these investigations have been discussed in previous sections, we shall consider in the present connection a number of interesting observations which have been made on the relations between the ionizing potentials of atoms from which one or more electrons have been removed.

Table 14, taken with slight modification from a recent paper by A. A. Noyes and A. O. Beckman (29.4), gives the relative values of the energy of ionization for the neutral stripped atoms of the first two series and part of the third in the periodic arrangement. The energy is given in terms of ν/R , where ν represents the wave number of the lowest level, and is therefore expressed as a *multiple of that required to ionize the H atom*, i.e., 13.54 volts ($= 109,678 \text{ cm.}^{-1}$; $= 312,368 \text{ calories per gram atom}$).

MOSELEY'S LAW FOR IONIZING POTENTIALS

The most striking feature about these data is the observation that if we plot the square roots (that is, $\sqrt{\nu/R}$ or $\sqrt{V_i}$) of the values corresponding to stripped atoms which are iso-electronic, against N , approximately straight lines are obtained. Figure 10 shows a series of such lines plotted for the series Li to Ne in all the states of ionization. Values for O_{VI} and F_{VII} have been obtained from published data by Millikan and Bowen. The regularity of the lines furnishes a beautiful illustration of the validity of Moseley's law for isoelectronic systems. As a matter of fact, the Moseley diagram in figure 7 for the $3s$ level in the series Na_I to S_{VI} represents a plot of the same data, since in the case of these one-electron systems this level is the lowest, and therefore corresponds to the ionizing energy. Furthermore, a comparison with the slope of

the line for which $\frac{\Delta\sqrt{\nu/R}}{\Delta N} = 1/2$, is evidence, as in the case of the K levels, for elements of high atomic number, that the total quantum number for these levels is equal to 2.

In table 14 a stair-like line has been drawn to the right of those values which correspond to the complete removal of the electrons in the valence shell. Thus in the case of Be it requires 0.70 units to remove the first s electron, 1.34 more units to remove the second (total = 2.04 to remove both valence electrons), but as the third electron belongs to the He "shell," the energy required to remove this one is considerably greater (11.4 units) and to remove the fourth one requires 16.0 units more. A better conception of the actual amounts of energy required for these various processes is obtained by converting these results into gram calories per gram atom.

Energy in kgm. calories per gram atom for removal of electrons in Be

ELECTRONS REMOVED	ENERGY	TOTAL ENERGY
1	218.7	218.7
2	418.6	637.3
3	3561	4198
4	4998	9198

The last column gives the *total energy* required for the removal of 1, 2, 3 or 4 electrons respectively. Comparing with these values the chemical energy of formation of BeCl_2 , 155 kgm. cal. per gram mol, it is seen that in general the energy of removal of one electron from the valence shell is of the same order of magnitude as chemical heats of formation, while that required for the removal of two or more electrons is usually larger than any chemical energy of formation. But the energy required to remove electrons from the rare gas shells underneath the valence shells is of a much different order of magnitude.

Now it is very interesting that the Moseley line shown in figure 10 for the series He_I , Li_{II} , Be_{III} , B_{IV} , and C_V (in all of which the electron removed belongs to the He shell) has a slope which is exactly twice that for the series Li_I to O_{VI} . (In the diagram the

slopes actually appear the same because the values of $(1/2)\sqrt{\nu/R}$ have been plotted.) This again is in agreement with the conclusion that in these atoms the electrons are in the 1s level.

Passing to the series Na-A, we find relations between the different ionizing energies similar to those observed for the first series. The Moseley diagrams for iso-electronic systems are shown in figure 11. The slope is not as well defined, but is

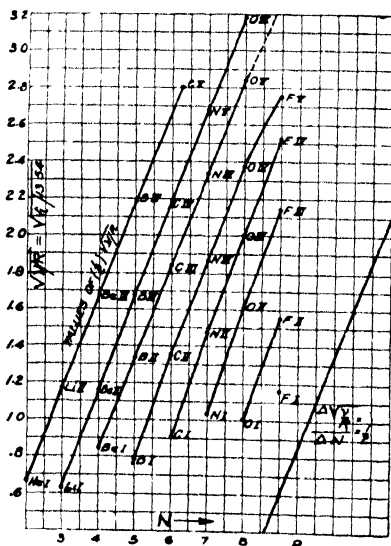


FIG. 10

FIG. 10. MOSELEY DIAGRAM OF IONIZATION POTENTIALS OF "STRIPPED" ATOMS OF SERIES Li TO F

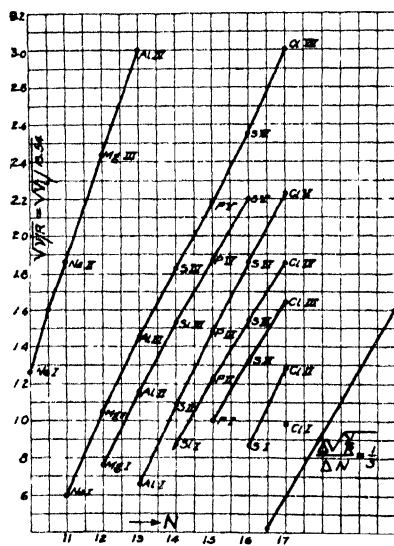


FIG. 11

FIG. 11. MOSELEY DIAGRAM OF IONIZATION POTENTIALS OF "STRIPPED" ATOMS OF SERIES Na TO Cl

approximately $1/3$, and here it is also observed that the values for the series Ne_I to Al_{IV} (which is analogous to the series He_I to C_V) shows the slope $1/2$, indicating that the electrons in these systems are in levels for which the total quantum number is 2, while in the systems $Na_I - Cl_{VII}$, etc., the electrons are in 3s levels.

For any one atomic system the value of the total energy for the removal of n electrons varies approximately linearly with n^2

(at least for larger values of n). Thus in the case of Ti, the values of n^2 and the corresponding values of the total energy (in the units used in table 14) are as follows:

n^2	1	4	9	16
Total energy	0 50	1 50	3 54	6 72
Ratio	0 50	0.38	0 39	0.42

There are other interesting relations which may be deduced from the data in table 14 but as a result of the discussion already given in this and previous sections it is evident that from a study of the energy terms in the spectra of stripped atoms it is possible to derive conclusions regarding the processes by which atoms are formed by the addition of successive electrons. We thus see in the spectra of these ionized atoms a process which is the reverse of that observed in x-ray spectra. But it is equally worthy of note that both lines of evidence lead to similar conclusions regarding the nature and distribution of electrons in the various atoms.

MAGNETIC PROPERTIES IN RELATION TO SPECTRAL TYPE

In view of the effect of magnetic fields on spectral terms it is natural to look for some relation between magnetic properties of atoms and the corresponding electron distribution. While it is not intended in the present connection to discuss this topic at great length, there are certain special points which may be mentioned briefly since they bear a direct relation to the conclusions based on investigations of spectral terms.

The fact that $\Sigma m = 0$ for all the electrons in a completed level is obviously in agreement with the observed diamagnetic properties of the rare gas elements. Also the experiments of Gerlach and Stern on the deflections of atomic beams in magnetic fields have served to throw additional light on the values of m for the atoms tested in their experiments. Finally Sommerfeld, Hund and Bose have attempted to develop theories of ionic paramagnetism on the basis of the electronic distributions assigned to the atoms in table 6.

In order to understand these theoretical considerations it is necessary, however, to refer briefly to the discussion of the Zeeman effect given previously. According to Larmor's theory, the magnetic moment due to an angular momentum of quantum number l is given by $\frac{e}{2m_0c} \cdot \frac{h}{2\pi} l$. The unit magnetic moment, $\frac{eh}{4\pi m_0c} = \mu_B$, is known as a Bohr *magneton* and corresponds to a magnetic moment per gram atom of $M_B = 6.06 \times 10^{23} \cdot \mu_B = 5584$ gauss. For an atomic system of quantum number, j , the magnetic moment, μ , is given by $\mu = \mu_B \cdot jg$ where g is Landé's factor. As mentioned previously, g gives the difference in energy between the different orientations in a magnetic field, and depends upon the values of L and j for the particular atomic state. Thus for a 2S_1 term in a weak magnetic field, $g = 2$, and hence $gj = 1$. Consequently the atoms of group I in the normal state possess one Bohr magneton.

MAGNETIC DEVIATION OF ATOMIC RAYS

Confirmation of this conclusion regarding the magnetic moments of the alkali atoms has been obtained by Gerlach and Stern in experiments which they have carried out at Frankfurt since 1921. As Stoner has stated (10):

In these investigations possible magnetic moments of atoms in the field direction are determined; the beauty of the experiments themselves, and the theoretical significance of the results, combine to make the research one of outstanding interest and fundamental importance.

In the experiments the element under investigation is heated in a small oven, and the evaporating atoms escape through an aperture; by means of a suitable system of slits a narrow unidirectional stream of atoms is produced, the velocity of the atoms depending on the oven temperature. The stream passes through a region in which a non-homogeneous (magnetic) field may be applied, the field and its gradient being at right angles to the stream, and the atoms impinge on a glass plate, forming a thin deposit, which can be rendered visible; from the traces obtained (giving the cross-section of the stream) with and without field the deviation produced by the field may be determined.

In the case of an atom for which the magnetic moment is one magneton two spots should be obtained in a magnetic field at equal distances from the zero position, corresponding to the two possible orientations of the magneton in a magnetic field, and actually this result was obtained for Cu, Ag and Au for each of which as shown in table 6 the normal term is 2S_1 .

No effect was observed for the vapors of Zn, Cd and Hg, which again is in accordance with the predictions based on the type of the normal term (1S_0), since for $j = 0$, the atom possesses no magnetic moment. Similar results were obtained for Sn and Pb for which $j = 0$. In the case of Ni and Fe, the results obtained by Gerlach and Stern are not so readily interpreted but indicate magnetic moments which are not those expected from the j -values for the normal terms.

PARAMAGNETISM OF IONS IN RELATION TO SPECTRAL TYPE

For an ionized atom the number of Bohr magnetons may be calculated from the electron distribution in the same manner as for the neutral atoms, that is, by the relation $\mu = \mu_B g j$ where j is the inner quantum number for the normal term in the ionized atom. Thus for Ca^+ and Sc^{++} the magnetic moment is one magneton, the same as for K . In the case of other elements of the second long period the first stage in the process of ionization consists in the removal of the two $4s$ electrons. Hence the normal terms for Ti^{++} and Fe^{++} are 3F_2 and 6D_4 respectively, while for Fe^{+++} , for which the electron configuration is $(3d)^5$, the normal term is 6S_5 .

Assuming that the same electron configurations are valid for the ions in solution it should be possible as was first pointed out by Sommerfeld (30), to calculate the magnetic moments from the j -values of the corresponding spectral terms. Hund (31) has applied these considerations to the calculation of the magnetic moments of the trivalent ions of the rare earths in solution and finds that the theoretical results are in very good agreement with the experimentally observed magnetic susceptibilities.

On the other hand, in the case of divalent and trivalent ions of the series Sc-Zn, the agreement is not so good. In explanation

it has been suggested by Hund that for these ions the different terms corresponding to any given electron configuration have energy values that are so near each other that probably the discrepancy is due to lack of knowledge of the particular value of j which should be used in the calculation of μ . It is true that Sommerfeld obtained good agreement by assuming that the ions are in an S state, for which, independently of the value of r , the number of magnetons is 1. However, this assumption cannot be justified on the basis of spectral terms.

In a recent paper, D. M. Bose (32) has brought forward the suggestion that in the case of the paramagnetic ions of the iron group only the magnetic moment due to the spin of the electrons is effective. This is, at first glance, the same hypothesis as that of Sommerfeld's since it corresponds to the assumption that the spectral term is of type $L = 0$ (that is, an S term). According to Bose, however, the difference between the two points of view is as follows:

According to Sommerfeld the magnetic moment of an atom or ion is an unalterable magnitude, $\mu = n \cdot \mu_B$ (where n = number of Bohr magnetons) and in a magnetic field the atom can orient itself as a whole along certain discrete directions with respect to the field. . . .

According to the view here presented (Bose's), μ is no longer a fixed magnitude and there is no definite orientation for the atom as a whole in a magnetic field. The magnetic moment is made up of the algebraic sum of the spin moments of the individual electrons in an uncompleted shell, some of which are oriented parallel and the rest anti-parallel to the direction of the field.

The agreement between the calculated values based on this theory and experimental observations is found to be satisfactory for trivalent and divalent ions of the elements Ti to Cu. But when the same theory is applied by Bose to the case of the trivalent ions of the rare earths, no agreement at all is found, whereas Hund's calculations, as mentioned previously, led to concordant results.

On the whole, therefore, we must regard this problem of correlation between magnetic moments and electron configurations as still in a developmental stage, and further investigations of a

theoretical nature are necessary before all the observed phenomena can be accounted for satisfactorily.

SPECTROSCOPIC TERMS AND VALENCE OF ATOMS

Naturally the chemist will seek for some connection between electronic configuration in the valence shell (or what is equivalent to this, viz., the spectral term) and valence bonds. In the light of present knowledge of atomic structure the essential elements in the conclusions deduced by Langmuir and Lewis over ten years ago regarding the mode of formation of polar and non-polar compounds still remain valid. In fact Langmuir's suggestions regarding electron distribution in the molecules of N_2 , CO and NO have received remarkable confirmation in the results obtained by R. S. Mulliken from an examination of energy levels in the band spectra of these molecules.

However, a number of papers have appeared in the past few years, in which attempts have been made to correlate more intimately the valence and co-valence relations of atoms with the electronic configurations as given in table 6. While it is intended to discuss the views of these writers at greater length in another paper, it may not be out of place to mention rather briefly some of these in the present connection.

Thus H. A. Grimm and A. Sommerfeld (33) have pointed out that the valence is intimately connected with the electronic configuration in the incompleted level, but their remarks are really a re-statement, in terms of Stoner's classification, of relations which Langmuir had derived on the basis of the octet theory.

In a previous section mention was made of Stoner's suggestion that the p , d and f levels really consist each of two sub-groups, which may be designated in accordance with the corresponding j values as p_1 , p_2 ; d_2 , d_3 , f_3 and f_4 respectively. The principal argument for this sub-division is the existence of corresponding x-ray levels as indicated in table 7 and figure 4. Stoner assumed that in the gradual completion of, say, the p level, the p_1 level would be completed first with 2 electrons and then the p_2 level with 4 electrons. Similarly the d_2 and d_3 sub-levels require 4 and 6 electrons respectively to complete them, while the f_3 and f_4 sub-levels require 6 and 8 electrons respectively.

In a series of papers published during the past two years, H. Lessheim and R. Samuel (34) have suggested that the valence relations of atoms are governed by the distribution of the valence electrons in these sub-levels. In arriving at this distribution they have modified Hund's method for deriving spectral terms in this respect that they add vectorially the angular momenta corresponding to the quantum numbers $m = j, j-1, \dots -j$, and $\mu = l - j$, for the case $m \leq l-s$, or $\mu = m - (j-s)$ for the case $m \geq l-s$. As a result they obtain for the normal term, the same result as Hund for the atoms in the first half of any level, but obtain the term with *lowest value* of j for the atoms in the second half (instead of highest value of j which Hund derives).

The method thus involves as a preliminary step the assignment of j values to the different electrons, that is, a division of the electrons in any one level into sub-levels with corresponding values of j . For instance, according to this point of view the distribution of electrons in the two p -levels for the atoms B to Ne (and similarly for the atoms Al to A) is as follows:

	p_1	p_2	NORMAL TERM]	NORMAL TERM (HUND)
B	1		$^2P_{1/2}$	$^2P_{1/2}$
C	2		3P_0	3P_0
N	3		$^4S_{3/2}$	$^4S_{3/2}$
O		4	3P_0	3P_2
F	1	4	$^2P_{1/2}$	$^2P_{3/2}$
Ne	2	4	1S_0	1S_0

The last two columns give the normal terms as derived by Lessheim and Samuel and compared with these, those derived by Hund. The fact that the latter is in agreement with spectroscopic observation would argue strongly against the view suggested by the former. Furthermore, it would appear meaningless to assign a j value to any one electron in a group of electrons since physically this quantum number can only apply to the atom as a whole and, therefore, to the resultant of the l 's and s 's for the individual electrons.

Obviously, with a distribution of electrons in two subgroups it

is easier to correlate these with observed valences. For example, in the case of S, $(s)^2(p)^4$, valences of 2, 4 and 6 are readily explained, while in that of Cl, for which the configuration is $(s)^2(p_1)(p_2)^4$, valences of 1, 3, 4, 5 and 7 are to be expected.

However, it would seem to the present writer that the arguments for assigning these distributions in sub-levels are not very sound, and a great many of the conclusions derived by Lessheim and Samuel can be derived without any recourse to assumptions such as they have made.

The problem of co-valency has been treated from the point of view of Bohr's earlier classification of electron levels by R. H. Fowler (35) and N. V. Sidgwick (36). The former suggests that in coördination compounds of Co, Fe, Pt, etc., we may have 12 electrons shared just as in Langmuir's theory, 8 electrons are shared. This would account for a co-valency of 6. Similarly Sidgwick concludes that in the acetyl acetone compounds of elements of higher atomic weight, a co-valency of 8 exists because of the tendency to share 16 electrons.

CONCLUDING REMARKS

While the details regarding spectral terms of a large number of atoms still remain to be investigated, the framework into which these details must fit has been constructed. We can now obtain a perspective, as it were, not only of the manner in which atoms are built up by the addition of successive electrons, but also of interactions between these electrons, compared to which the "harmony of the celestial spheres" is a very simple phenomenon.

Once more the periodic arrangement appears as the most significant guiding principle in the whole of physical science; but where Mendeljeff and Newman saw only certain chemical aspects and atomic weights, the work of the spectroscopists, interpreted along the lines first suggested by Bohr, has led to a revelation that beyond the gross properties of atoms there exist as governing factors a nuclear charge and electrons in definite configurations.

The next great advance in chemistry is going to consist in developing the relations between the energy levels in atoms and those in the molecules which result from their combination. A

beginning has already been made by investigators like R. S. Mulliken, R. T. Birge, J. Franck and others, in the study of energy levels in molecules from band spectra and the correlation of heats of dissociation with observations on absorption limits in these band spectra. But a great deal remains to be accomplished and more and more the physical chemist will find in the investigations of the spectroscopic phenomena of molecules guidance to the ultimate understanding of the hitherto illusive conception of chemical bonds.

In conclusion, the writer wishes to express his appreciation of the help which he has received towards an understanding of Hund's ideas from Professor Russell's published papers and personal correspondence with him on some of the difficulties.

REFERENCES

- (1) N. BOHR: *The Theory of Spectra and Atomic Constitution*, Cambridge University Press 1922.
- (2) K. T. COMPTON AND F. L. MOHLER: *Critical Potentials*, Bull. Nat. Research Council, Vol. 9, Part 1 (1924).
- (3) A. FOWLER: *Report on Series in Line Spectra*, London (1922).
- (4) F. HUND: *Linienpektren und periodisches System der Elemente*, Julius Springer, Berlin (1927).
- (5) A. SOMMERFELD. *Atombau u. Spektrallinien* 4 Auflage. F. Vieweg and Sohn, Braunschweig (1924)
- (6) P. D. FOOTE AND F. L. MOHLER *The Origin of Spectra*, Chem. Catalog Co., New York (1922)
- (7) P. D. FOOTE *J. Franklin Inst.*, **198**, 344 (1924).
- (8) H. N. RUSSELL AND F. A. SAUNDERS: *Astrophys. J.*, **61**, 38 (1925).
- (9) E. BACK AND A. LANDÉ: *Zeemaneffekt und Multiplettstruktur der Spektrallinien*, J. Springer, Berlin (1925)
- (10) E. C. STONER. *Magnetism and Atomic Structure*. E. P. Dutton and Company, New York (1926).
- (11) G. E. UHLENBECK AND S. GOUDSMIT: *Naturwissenschaften*, **13**, 953 (1925); *Nature*, **117**, 264 (1926).
- (12) E. C. STONER: *Phil. Mag.*, **48**, 719 (1924).
- (13) J. D. MAIN-SMITH *Chemistry and Atomic Structure*, D. Van Nostrand Company, New York (1924).
- (14) N. BOHR AND COSTER: *Z. Physik.*, **12**, 342 (1923).
- (15) J. S. McLENNAN, A. B. McLAY AND H. G. SMITH: *Proc. Roy. Soc. London [A]*, **112**, 76 (1926).
- (16) F. HUND: *Z. Physik*, **33**, 345 (1925).
- (17) W. PAULI: *Z. Physik*, **31**, 765 (1925).
- (18) A. C. HAUSMANN. *Astrophys. J.*, **66**, 333 (1927).

- (19) A. FOWLER: Proc. Roy. Soc. London, [A] **103**, 722 (1923).
- (20) I. S. BOWEN AND R. A. MILLIKAN: Phys. Rev., **25**, 295 (1925).
- (21) A. FOWLER: Nature, 223 (1924).
- (22) Further references to papers by Millikan and Bowen are as follows:
 - (1) Phys. Rev., **23**, 1 (1924). Gives the original data on extreme ultra-violet spectra of a number of elements.
 - (2) Phys. Rev., **26**, 310 (1925). Series spectra of two valence electrons of boron and carbon.
 - (3) Phys. Rev., **27**, 144 (1926). Stripped Oxygen, O_{VI} , the pp^1 Group in O_V .
 - (4) Phys. Rev., **28**, 256 (1926). Series spectra of B_{cI} and Be_{cII} .
 - (5) Phys. Rev., **29**, 231 (1927). Series spectra of B , C , N , O and F .
 - (6) Phys. Rev., **29**, 510 (1927). Series spectra of Ionized Phosphorus, P_{II} .
- (23) H. N. RUSSELL: Astrophys. J., **66**, 347 (1927).
- (24) R. A. MILLIKAN AND I. S. BOWEN: Phys. Rev., **24**, 209 (1924).
- (25) W. HEISENBERG AND P. JORDAN: Z. Physik., **37**, 263 (1926).
- (26) R. C. GIBBS AND H. E. WHITE: Proc. Nat. Acad. Sci., **12**, 448 (1926); *ib.*, **12**, 551 (1926).
- (27) R. C. GIBBS AND H. E. WHITE: Proc. Nat. Acad. Sci., **13**, 525 (1927).
- (28) H. N. RUSSELL: Astrophys. J., **66**, 233 (1927).
- (29) References on ionizing potentials:
 - (1) Reference (2) Values for H, Li, Na, K, Rb, Cs, Cu, Ag, Au; Mg, Ca, Sr, Ba, Zn, Cd, Hg; Al, Ga, In, Tl; Pb; O, S, Se; Mo; He, Ne, A, Kr, Xe.
 - (2) Reference (28) Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu.
 - (3) GREEN AND LORING: Phys. Rev., **30**, 574 (1927) Sr.
 - (4) A. A. NOYES AND A. O. BECKMAN: Proc. Nat. Acad. Sci., **13**, 737 (1927). Summarizes data published by Millikan and Bowen; C, Be, Si, P, Cl.
 - (5) H. DINGLE: Proc. Roy. Soc., **117**, 407 (1927). Fluorine. V_i is given as 18.6 volts, which is slightly higher than that given in 29.4.
 - (6) L. A. SOMMER: Z. Physik., **45**, 147 (1927) Ru, Rh, Pd.
 - (7) A. G. SHENSTONE: Phys. Rev., **29**, 380 (1927) Ni, Pd.
- (30) A. SOMMERFELD: Z. Physik., **19**, 221 (1923); also ATOMBAU: pp. 630-648.
- (31) F. HUND: Z. Physik., **33**, 855 (1925); also Linienspektren, p. 179.
- (32) D. M. BOSE: Z. Physik., **43**, 864 (1927).
- (33) H. A. GRIMM AND A. SOMMERFELD: Z. Physik., **36**, 36 (1926).
- (34) H. LESSHEIM AND R. SAMUEL: Z. Physik., **40**, 220 (1926), *ib.*, **42**, 614 (1927); also "Die Valenzzahl," Fortschritte der Chem. Physik. u. physikal. Chem., **19**, Heft 3 (1927).
- (35) R. H. FOWLER: Trans. Faraday Soc., **19**, 459 (1923).
- (36) N. V. SIDGWICK: Trans. Faraday Soc., **19**, 469 (1923).
- (37) A. SOMMERFELD: Atomic Structure, 3rd ed., N. Y., Dutton, (1923), p. 294.

THE APPLICATION OF THE QUANTUM MECHANICS TO THE STRUCTURE OF THE HYDROGEN MOLE- CULE AND HYDROGEN MOLECULE-ION AND TO RELATED PROBLEMS

LINUS PAULING

Gates Chemical Laboratory, California Institute of Technology

I. INTRODUCTION

Many attempts were made to derive with the old quantum theory structures for the hydrogen molecule, H_2 , and the hydrogen molecule-ion, H_2^+ , in agreement with the experimentally observed properties of these substances, in particular their energy contents. These were all unsuccessful, as were similar attempts to derive a satisfactory structure for the helium atom. It became increasingly evident that in these cases the straightforward application of the old quantum theory led to results definitely incompatible with the observed properties of the substances, and that the introduction of variations in the quantum rules was not sufficient to remove the disagreement. (For a summary of these applications see, for example, Van Vleck (1).) This fact was one of those which led to the rejection of the old quantum theory and the origination of the new quantum mechanics. The fundamental principles of the quantum mechanics were proposed by Heisenberg (2) in 1925. The introduction of the matrix algebra (3) led to rapid developments. Many applications of the theory were made, and in every case there was found agreement with experiment. Then the wave equation was discovered by Schrödinger (4), who developed and applied his wave mechanics independently of the previous work. Schrödinger's methods are often considerably simpler than matrix methods of calculation, and since it has been shown (5) that the wave mechanics and the matrix mechanics are mathematically identical, the wave equation is

generally used as the starting point in the consideration of the properties of atomic systems, in particular of stationary states.

The physical interpretation of the quantum mechanics and its generalization to include aperiodic phenomena have been the subject of papers by Dirac, Jordan, Heisenberg, and other authors. For our purpose, the calculation of the properties of molecules in stationary states and particularly in the normal state, the consideration of the Schrödinger wave equation alone suffices, and it will not be necessary to discuss the extended theory.

In the following pages, after the introductory consideration of the experimentally determined properties of the hydrogen molecule and molecule-ion, a unified treatment of the application of the quantum mechanics to the structure of these systems is presented. In the course of this treatment a critical discussion will be given the numerous and scattered pertinent publications. It will be seen that in every case the quantum mechanics in contradistinction to the old quantum theory leads to results in agreement with experiment within the limit of error of the calculation. It is of particular significance that the straightforward application of the quantum mechanics results in the unambiguous conclusion that two hydrogen atoms will form a molecule but that two helium atoms will not; for this distinction is characteristically chemical, and its clarification marks the genesis of the science of sub-atomic theoretical chemistry.

II. THE OBSERVED PROPERTIES OF THE HYDROGEN MOLECULE AND MOLECULE-ION

The properties of the hydrogen molecule and molecule-ion which are the most accurately determined and which have also been the subject of theoretical investigation are ionization potentials, heats of dissociation, frequencies of nuclear oscillation, and moments of inertia. The experimental values of all of these quantities are usually obtained from spectroscopic data; substantiation is in some cases provided by other experiments, such as thermochemical measurements, specific heats, etc. A review of the experimental values and comparison with some theoretical

results published by Birge (7) has been used as the basis for the following discussion.

The ultraviolet absorption spectrum of hydrogen was analyzed by Dieke and Hopfield (8). They identified the three lowest

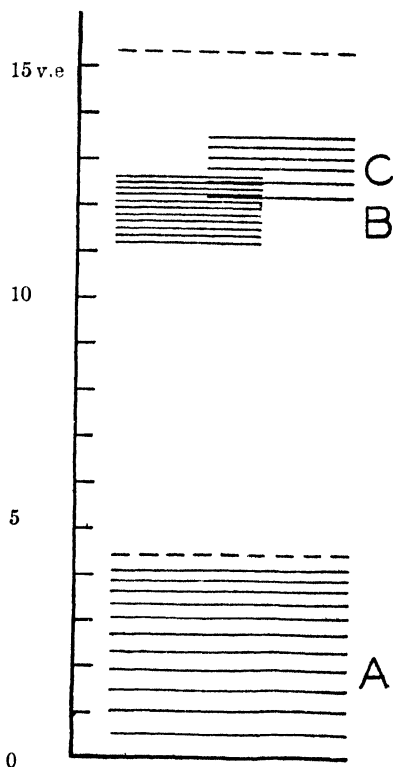


FIG 1. ENERGY LEVEL DIAGRAM FOR THE THREE LOWEST ELECTRONIC STATES OF THE HYDROGEN MOLECULE, SHOWING SUCCESSIVE OSCILLATIONAL LEVELS

electronic states, indicated in figure 1 by the symbols *A*, *B*, and *C*, and for each a number of states of oscillation of the nuclei, also shown in the figure. In addition there is a fine-structure of each oscillational state due to rotation of the molecule. A number of these levels were independently obtained by Witmer (9) from the analysis of the Lyman bands. Richardson (10) has analyzed

bands in the visible, and on the basis of comparison of the oscillational and rotational fine-structure has identified the lower states of some of these bands with Dieke and Hopfield's *B* and *C* states. The upper levels of Richardson's bands correspond to an electron in successive excited states, and by means of an assumed Rydberg formula Birge has carried out the extrapolation to ionization, and has obtained for the ionization potential of the hydrogen molecule the value

$$I_{H_2} = 15.34 \pm 0.01 \text{ volts}$$

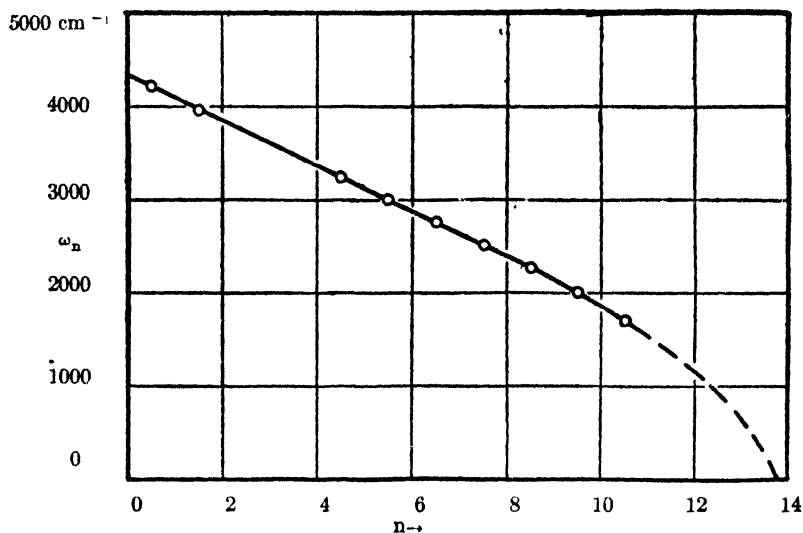


FIG. 2. THE OSCILLATIONAL FREQUENCY FOR THE NORMAL STATE OF THE HYDROGEN MOLECULE AS A FUNCTION OF THE OSCILLATIONAL QUANTUM NUMBER

The area under the curve gives the heat of dissociation

He also states that he has verified this value by means of the heat of dissociation of various excited states of the molecule as obtained by extrapolating the oscillational levels in the way described below.

Ionization by electron impact has been shown (11) to occur at about 16.1 volts. Condon has given the explanation of the discrepancy between this and the band spectrum value in terms of

a general phenomenon discovered and interpreted by Franck (12). We shall see later that the equilibrium internuclear distance for H_2 is about 0.75 Å, and for H_2^+ , 1.06 Å. When H_2 in the normal state is ionized by electron impact to H_2^+ the nuclei do not have time to move, but are left some distance from their new equilibrium positions. As a result large nuclear oscillations occur, estimated by Condon to correspond to about 1 volt-electron of oscillational energy; and this energy in addition to the ionization energy must be supplied by the impacting electron.

The best experimental value of the heat of dissociation of H_2 is that obtained by Witmer (9) by extrapolating the oscillational levels of the normal state of the molecule to dissociation. The restoring force acting on the two nuclei becomes smaller as the nuclei get farther apart, and as a result the oscillational frequency in successive oscillational states becomes smaller and smaller. For H_2 in the normal state this oscillational frequency

$$\omega_n = \frac{1}{h} \frac{\partial E_n}{\partial n}$$

is represented in figure 2 for values of n , the oscillational quantum number,¹ from 0 to 11. The curve was extrapolated by Witmer as shown by the dotted line; the area under it is equal to the heat of dissociation, and gives

$$D_{H_2} = 4.34 \pm 0.1 \text{ volt-electrons}$$

It is of interest to note that dissociation, represented in figure 1 by the dotted line, is very close to the highest observed oscillational-rotational state. Indeed the oscillational and rotational energy of the highest observed level was 4.10 v.e. (94,600 cal/mole), which must be a lower limit to the heat of dissociation.

The value 4.34 v.e. is equal to 100,000 cal/mole. Thermochemical measurements are in satisfactory agreement with this spectroscopic result. Thus Isnardi's experiments (13) on the thermal conductivity of partially dissociated hydrogen give, with the computational error discovered by Wohl (14) corrected, a

¹ The true oscillational quantum number has the values 1/2, 3/2, 5/2, ... For convenience we shall use in this paper the integers obtained by subtracting 1/2 from these values, unless specific mention is made of an alternative procedure.

value between 96,765 and 102,257 cal/mole. Langmuir's most recent value from his similar measurements (15) is 97,000 cal/mole.

From the rotational fine-structure of the A levels Hori (16) obtained the value

$$I_0 = 0.467 \times 10^{-40} \text{ g. cm.}^2$$

for the moment of inertia of the molecule in the normal state, corresponding to an equilibrium internuclear distance

$$r_0 = 0.76 \text{ \AA}$$

TABLE I
The properties of the hydrogen molecule

		E_e, cm^{-1}	n^*	ω_0, cm^{-1}	$r_0, \text{\AA}$
1	1 ¹ S(A)	0	0 9396	4264	0 76
2	2 ¹ S(B)	90083	1 7920	1325	1 55
3	2 ³ P	(94735)	(1 9281)	2390	0 97
4	2 ³ S	94906 7	1 9337	2593 82	1.08
5	2 ¹ P	(95469)	(1 9526)		
6	C	99086	2 0882	2380	1 06
7	3 ³ S	111427	2 9261		
8	3 ³ P	111518 1	2 9365	2306 94	1 136
9	3 ¹ P	111656 8	2 9526	2373 89	
10	4 ³ P	117169 9	3 9395	2276 45	1 145
11	4 ¹ P	117216 9	3 9526	2325 6(?)	
12	5 ³ P	119744 2	4 941	2251	1 168
13	6 ³ P	121130.2	5 942	2229 (?)	1 166
14	7 ³ P	121961 0	6 942		
15	8 ³ P	122498 3	7 942		
16	H ₂ ⁺	124237		2247	(1 06)

This result is independently verified by Dennison (17) who has recently given a satisfactory theory of the specific heat of hydrogen. The observed specific heat as interpreted by Dennison requires that I_0 be equal to $0.464 \times 10^{-40} \text{ g. cm.}^2$. The very recent measurements by Cornish and Eastman (18) of the specific heat of hydrogen from the velocity of sound are said to agree very well with Dennison's theory if I_0 be given the value of $0.475 \times 10^{-40} \text{ g. cm.}^2$.

The oscillational frequency of the nuclei in H₂ in the normal

state was calculated by Birge from the measurements of Witmer, Dieke and Hopfield, and Hori to be

$$\omega_{H_2} = 4264 \text{ cm.}^{-1}$$

By extrapolating the observed oscillational frequencies for various excited states of H_2 forming a part of a Rydberg series, Birge obtained for H_2^+ in the normal state the value

$$\omega_{H_2^+} = 2247 \text{ cm.}^{-1}$$

A summary of information regarding various electronic states of the hydrogen molecule is given in table 1, quoted from Birge.

The symbols in the second column represent the electronic state; in particular the first number is the total quantum number of the excited electron. We shall see later that in one case at least the symbol is probably incorrect. The third column gives the wave-number of the lowest oscillational-rotational level, the fourth the effective quantum number, the fifth and sixth the oscillational wave-number and the average internuclear distance for the lowest oscillational-rotational level. The data for H_2^+ were obtained by extrapolation, except r_0 , which is Burrau's theoretical value (Section VIa).

The interrelation of these quantities and comparison with theoretical results will be discussed in the following sections of this paper.

III. THE HYDROGEN ATOM

The wave equation representing a conservative Newtonian dynamical system is

$$\Delta\psi + \frac{8\pi^2}{h^2} (W - V(q_k)) \psi = 0 \quad (1)$$

with the conditions that ψ , the wave function or eigenfunction, be everywhere continuous, single-valued, and bounded. W and $V(q_k)$ are the energy constant and the potential energy; and the differential operations are with respect to coördinates whose line-element is given by

$$ds^2 = 2T(q_k, \dot{q}_k) dt^2,$$

in which T is the kinetic energy expressed as a function of the velocities. Only certain functions satisfy these requirements in any given case; to each there corresponds a characteristic value of the energy constant W . For a hydrogen-like atom with fixed nucleus of charge Ze the potential energy is $-\frac{Ze^2}{r}$, and the wave equation is

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{8\pi^2 m}{h^2} \left(W + \frac{Ze^2}{r} \right) \psi = 0 \quad (2)$$

On writing for the eigenfunction

$$\psi_{nlm} = X_{nl}(r) Y_{lm}(\vartheta) Z_m(\varphi), \quad (3)$$

the wave equation can be resolved into three total differential equations, with the solutions (4) (19)

$$\left. \begin{aligned} X_{nl}(r) &= \left\{ \left(\frac{2Z}{na_0} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^2} \right\}^{\frac{1}{2}} e^{-\frac{\xi}{2}} \xi^l L_{n+l}^{(2l+1)}(\xi) \\ \text{with} \quad \xi &= \frac{2Z}{na_0} r, \quad a_0 = \frac{h^2}{4\pi^2 me^2} = 0.529 \text{ \AA}. \\ Y_{lm}(\vartheta) &= \left\{ (l + \frac{1}{2}) \frac{(l-m)!}{(l+m)!} \right\}^{\frac{1}{2}} P_l^m(\cos \vartheta) \\ Z_m(\varphi) &= \frac{1}{\sqrt{2\pi}} e^{im\varphi} \end{aligned} \right\} \quad (4)$$

$L_{n+l}^{(2l+1)}(\xi)$ represents the $(2l+1)^{\text{th}}$ derivative of the $(n+l)^{\text{th}}$ Laguerre polynomial (20); and $P_l^m(\cos \vartheta)$ is Ferrers' associated Legendre function of the first kind, of degree l and order m . $Y_{lm} Z_m$ thus constitutes a tesseral harmonic (21). The ψ 's are in this form orthogonal and normalized, so that they fulfill the conditions

$$\int \psi_{nlm} \psi_{n'l'm'} d\Omega = \begin{cases} 1 & \text{for } n = n', l = l', m = m' \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

in which $d\Omega$ is the element of volume. The parameter n , the principal quantum number, can assume the values 1, 2, 3, . . . ;

l , the analogue of the azimuthal quantum number, the values 0, 1, 2, . . . $n-1$; and m , the magnetic quantum number, the values 0, ± 1 , ± 2 , . . . $\pm l$.

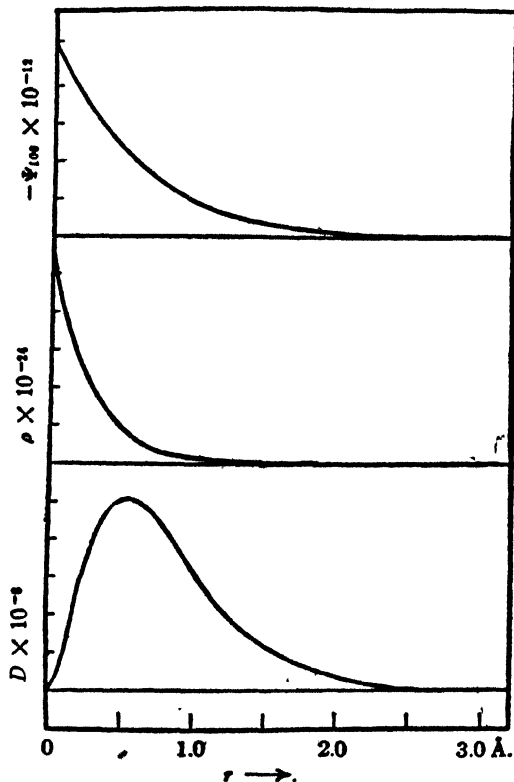


FIG. 3

The eigenfunction ψ_{100} , the electron density $\rho = \psi_{100}^2$, and the electron distribution function $D = 4 \pi r^2 \rho$ of the normal hydrogen atom as functions of the distance r from the nucleus.

The normal state of the atom is that with $n = 1$, $l = 0$, $m = 0$. The corresponding eigenfunction is

$$\psi_{100} = \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{Zr}{a_0}} \quad (6)$$

The quantity ψ^2 as a function of the coördinates is interpreted as the probability of the corresponding microscopic state of the system; in this case the probability that the electron occupies a certain position relative to the nucleus. It is seen from equation 6 that in the normal state the hydrogen atom is spherically symmetrical, for ψ_{100} is a function of r alone. The atom is furthermore not bounded, but extends to infinity; the major portion is, however, within a radius of about $2a_0$ or 1\AA . In figure 3 are represented the eigenfunction ψ_{100} , the average electron density $\rho = \psi_{100}^2$ and the radial electron distribution $D = 4\pi r^2\rho$ for the normal state of the hydrogen atom.

The energy values corresponding to the various stationary states are found from the wave equation to be those deduced originally by Bohr with the old quantum theory; namely,

$$W_{nlm} = -\frac{2\pi^2 m e^2 Z^2}{n^2 h^2} = -\frac{Z^2 e^2}{2n^2 a_0} \quad (7)$$

The energy of the normal state of hydrogen is

$$W_{II} = -\frac{e^2}{2a_0} = -13.54 \text{ v e} \quad (8)$$

I_H , the ionization potential of hydrogen, accordingly is equal to 13.54 volts.

This simplified treatment does not account for the fine-structure of the hydrogen spectrum. It has been shown by Dirac (22) that the assumption that the system conforma to the principles of the quantum mechanics and of the theory of relativity leads to results which are to a first approximation equivalent to attributing to each electron a spin; that is, a mechanical moment and a magnetic moment, and to assuming that the spin vector can take either one of two possible orientations in space. The existence of this spin of the electron had been previously deduced by Uhlenbeck and Goudsmit (23) from the empirical study of line spectra. This result is of particular importance for the problems of chemistry.

IV. THE PERTURBATION THEORY OF THE QUANTUM MECHANICS. THE RESONANCE PHENOMENON

The first-order perturbation theory of the quantum mechanics (4, III) is very simple when applied to a non-degenerate state of a system; that is, a state for which only one eigenfunction exists. The energy change W^1 resulting from a perturbation function f is just the quantum mechanics average of f for the state in question; i.e., it is

$$W^1 = \int f \psi^2 d\Omega \quad (9)$$

As an example we may calculate the energy of the helium atom in its normal state (24). Neglecting the interaction of the two electrons, each electron is in a hydrogen-like orbit, represented by equation 6; the eigenfunction of the whole atom is then ψ_{100} (1) ψ_{100} (2), where (1) and (2) signify the first and the second electron. The perturbation function is the electronic interaction $\frac{e^2}{r_{12}}$; and the perturbation energy is

$$W^1 = \iint \frac{e^2}{r_{12}} \psi_{100}^2 (1) \psi_{100}^2 (2) d\Omega_1 d\Omega_2 \quad (10)$$

in which subscripts refer to the two electrons. This integral has the value $\frac{5}{8} \frac{Ze^2}{a_0}$. The energy of the unperturbed system was (equation 7) — $2 \frac{Z^2 e^2}{2a_0} = -108.4$ v.e., giving a total energy of — $\frac{11}{4} \frac{e^2}{a_0}$ or —74.5 v.e. The experimentally determined value is —78.8 v.e. Thus the first-order approximation reduces the discrepancy from 29.6 v.e. to —4.3 v.e. A more accurate theoretical calculation (25) has led to —77.9 v.e.

If the unperturbed system is degenerate, so that several linearly independent eigenfunctions correspond to the same energy value, then a more complicated procedure must be followed. There can always be found a set of eigenfunctions (the zeroth order eigenfunctions) such that for each the perturbation energy is given by equation 9; and the perturbation theory provides the

method for finding these eigenfunctions (4, III). This result is of importance for systems containing two or more electrons. Let us consider, for example, an excited helium atom. Neglecting the interaction of the two nuclei the system may be described by saying that the electrons are in two different orbits, represented by the hydrogen-like eigenfunctions ψ and φ , say. The eigenfunction for the atom is, then, $\psi(1)\varphi(2)$. But on interchanging the electrons to give $\psi(2)\varphi(1)$ the energy of the system is unchanged, so that the system is degenerate. The perturbation theory leads to the result that the correct eigenfunctions are not $\psi(1)\varphi(2)$ and $\varphi(1)\psi(2)$, but rather

$$\Psi_S = \frac{1}{\sqrt{2}} \{ \psi(1)\varphi(2) + \varphi(1)\psi(2) \} \quad (11a)$$

and

$$\Psi_A = \frac{1}{\sqrt{2}} \{ \psi(1)\varphi(2) - \varphi(1)\psi(2) \} \quad (11b)$$

Ψ_S is said to be symmetric in the coördinates of the two electrons, for interchanging them leaves the eigenfunction unchanged, while Ψ_A is antisymmetric, for interchanging the electrons changes the sign of the eigenfunction.

Substitution of these eigenfunctions in equation 9 leads to the result

$$\begin{aligned} W^1_S &= H_{11} + H_{12} \\ W^1_A &= H_{11} - H_{12} \end{aligned} \quad (12)$$

with

$$\begin{aligned} H_{11} &= \int \int \psi^2(1) \varphi^2(2) d\Omega_1 d\Omega_2 \\ H_{12} &= \int \int \psi(1) \varphi(1) \psi(2) \varphi(2) d\Omega_1 d\Omega_2. \end{aligned}$$

If the electrons did not change positions; that is, if $\psi(1)\varphi(2)$ or $\varphi(1)\psi(2)$ were the correct eigenfunction, the perturbation energy would be H_{11} alone. The added or subtracted term H_{12} results from one electron jumping from one orbit (ψ) to the other (φ) at the same time that the other electron makes the reverse jump. For this reason $\pm H_{12}$ is called the interchange or resonance energy. This phenomenon, called the interchange or resonance phenomenon, was discovered by Heisenberg and Dirac (26).

There is no classical analogue of it save the trivial case of the resonance of two similar harmonic oscillators.

The interchange energy of electrons is in general the energy of the non-polar or shared-electron chemical bond.

V. THE PROPERTIES OF MOLECULES ACCORDING TO THE QUANTUM MECHANICS

The procedure to be followed in the theoretical discussion of the structure of molecules has been given by Born and Oppenheimer (27), who applied the perturbation theory (to the fourth order) to a system of nuclei and electrons. They showed that the electronic energy is first to be calculated for various arrangements of the nuclei fixed in space. The stable state will then be that for which the so-calculated electronic energy added to the inter-nuclear energy is a minimum. The nuclei will then undergo oscillations about their equilibrium positions, with the electronic and nuclear energy as the restoring potential; and the molecule as a whole will undergo rotations about axes passing through its center of mass.

The justification was also given for the assumption made originally by Franck (12) that during an electron transition the nuclei retain the configuration characteristic of the initial state.

These results were true for classical mechanics and the old quantum theory, and had been assumed without proof by many people before the work of Born and Oppenheimer was published.

VI. THE HYDROGEN MOLECULE-ION

a. Numerical solution of the wave equation

The system to be considered consists of two nuclei and one electron. For generality let the nuclear charges be $Z_A e$ and $Z_B e$. From Born and Oppenheimer's results it is seen that the first step in the determination of the stationary states of the system is the evaluation of the electronic energy with the nuclei fixed an arbitrary distance apart. The wave equation is

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{8\pi^2 m}{h^2} \left(W + \frac{Z_A e^2}{r_A} + \frac{Z_B e^2}{r_B} \right) \psi = 0 \quad (13)$$

in which r_A and r_B represent the distance of the electron from nucleus A and nucleus B respectively. If elliptic coordinates ξ , η , and φ , defined by the equations

$$\left. \begin{aligned} \xi &= \frac{r_A + r_B}{r_{AB}} \\ \eta &= \frac{r_A - r_B}{r_{AB}} \\ \varphi &= \cos^{-1} \frac{x}{\sqrt{x^2 + y^2}} \end{aligned} \right\} (14)$$

in which r_{AB} is the internuclear distance, be introduced, the partial differential equation becomes separable into three total differential equations. For introducing

$$\psi = \Xi(\xi) H(\eta) \Phi(\varphi), \quad (15)$$

it reduces to

$$\frac{d^2 \Phi}{d\varphi^2} = m^2 \Phi \quad (16a)$$

$$\frac{d}{d\xi} \left[(\xi^2 - 1) \frac{d\Xi}{d\xi} \right] + \left(-\lambda \xi^2 + 2\rho \xi + \frac{m^2}{\xi^2 - 1} + \mu \right) \Xi = 0 \quad (16b)$$

$$\text{with } 1 \leq \xi \leq \infty$$

and

$$\frac{d}{d\eta} \left[(1 - \eta^2) \frac{dH}{d\eta} \right] + \left(\lambda \eta^2 - 2b\eta + \frac{m^2}{1 - \eta^2} - \mu \right) H = 0 \quad (16c)$$

$$\text{with } -1 \leq \eta \leq +1$$

in which

$$\lambda = -\frac{2\pi^2 m r_{AB}^2 W}{h^2} = \rho^2 \frac{W}{4W_1},$$

$$\rho = \frac{r_{AB} (Z_A + Z_B)}{2a_0}$$

$$b = \frac{r_{AB} (Z_A - Z_B)}{2a_0}$$

and m , λ , and μ are undetermined parameters. The solutions of Equation 16a are known; in order that Φ be an eigenfunction m must assume only the values 0, ± 1 , ± 2 , . . . Similarly in order for Ξ and H to be eigenfunctions λ and μ must assume specific values, with the accompanying determination of the energy constant W as a function of the nuclear separation r_{AB} .

Many efforts have been made to solve these equations analytically, but so far they have all been unsuccessful, and little has been published regarding them. Some unsatisfactory work has been reported by Alexandrow (28). Very recently a short report of a paper read by Wilson before the Royal Society has appeared (29). It is probable, in view of the vigor with which it is being attacked, that the problem will be solved completely before very long.

The problem has already been solved for the normal state of the hydrogen molecule-ion ($Z_A = Z_B = 1$) by the use of numerical methods. A rather complete account of these calculations of Burrau (30) will be given here, since the journal in which they were published is often not available.

The numerical solution of the equations was carried out in the following way. For the lowest state m is equal to 0, and for hydrogen, with $Z_A = Z_B = 1$, $b = 0$. Introducing new variables $\sigma_\xi = -\frac{1}{\Xi} \frac{d\Xi}{d\xi}$ and $\sigma_\eta = -\frac{1}{H} \frac{dH}{d\eta}$, Equations 16b and 16c become

$$\frac{d\sigma_\xi}{d\xi} - \sigma_\xi^2 - \frac{\lambda \xi^2 - 2(\rho - \sigma_\xi)\xi - \mu}{\xi^2 - 1} \quad (16b')$$

and

$$\frac{d\sigma_\eta}{d\eta} = \sigma_\eta^2 - \frac{\lambda \eta^2 + 2\sigma_\eta \eta - \mu}{1 - \eta^2} \quad (16c')$$

For a given value of λ , σ_η is expanded as a power series satisfying Equation 16c' about the points $\eta = 0$ and $\eta = 1$ (or -1). It is found that these series coincide at $\eta = 1/2$ only when μ has a certain value. In this way a relation between λ and μ is found.

For each value of μ a similar treatment is given Equation 16b', resulting in the determination of the corresponding value of ρ , and hence of the relation between W and ρ . The results of the calculation are given in table 2, taken directly from Burrau. The electronic energy is given in the third column. To it must be added the internuclear energy $\frac{e^2}{r_{AB}} = -\frac{^2W_H}{\rho}$ (fourth column) to obtain the total energy given in the fifth column.

The relation between W and ρ is shown in figure 4; that for W_{total} in figure 5. It is seen that the equilibrium distance of the nuclei is $\rho = 2.0 \pm 0.1$, or $r_0 = 1.06 \pm 0.05 \text{ \AA}$. The corresponding total energy is $W_{\text{total}} = 1.204 \pm 0.002 W_H$ or -16.30

TABLE 2
Energy of the hydrogen molecule-ion

ρ	λ	$\frac{W}{W_H} \approx \frac{4\lambda}{\rho^2}$	$\frac{2}{\rho}$	$\frac{W_{\text{total}}}{W_H}$
0.0	0 000	4.000	∞	∞
1 0	0.724	2.896	2.000	0.896
1.3	1 119	2.648	1.538	1.110
1 6	1 559	2.436	1.250	1.186
1 8	1 870	2.309	1.111	1.198
2 0	2 204	2.204	1.000	1.204
2 2	2 552	2.109	0.909	1.200
2 4	2 917	2.025	0.833	1.192
2.95	3 995	1.836	0.678	1.158
∞		1.000	0.000	1.000

± 0.03 v.e. To this there is to be added the oscillational energy $\frac{1}{2}h\omega_0$ of the lowest oscillational state. The frequency ω_0 is obtained from the curvature of the W_{total} curve, and leads to $\frac{1}{2}h\omega_0 = 0.14$ v.e. according to Condon, who corrected an error made by Burrau (Condon's value is reported by Birge (7)). The energy of H_2^+ in the normal state is thus

$$W_{H_2^+} = -16.16 \pm 0.03 \text{ v.e.}$$

It is now possible to check the relation

$$I_{H_2} + I_{H_2^+} = D'_{H_2} + 2 I_H,$$

in which D'_{H_2} is the dissociation potential of hydrogen. With $I_{H_2} = + 15.34$ v. (Section 2), and $I_H = + 13.54$ v. This equation leads to

$$D_{H_2} = 4.42 \text{ v.e.} = 102000 \text{ cal/mol.}$$

for the energy of dissociation of the hydrogen molecule. This is within the limit of error equal to the band spectrum value 4.34

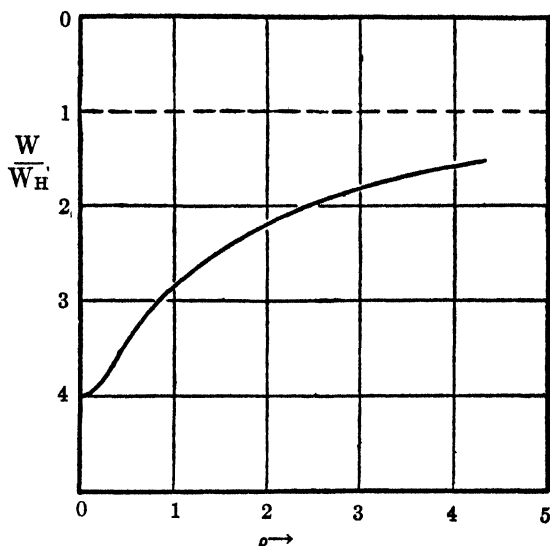


FIG. 4. THE ELECTRONIC ENERGY OF THE HYDROGEN MOLECULE-ION IN THE NORMAL STATE AS A FUNCTION OF THE DISTANCE BETWEEN THE TWO NUCLEI (BURRAU)

± 0.10 v.e., and is to be accepted as the most accurate determination of the heat of dissociation of hydrogen.

The heat of dissociation of H_2^+ into H and H^+ is

$$D_{H_2^+} = W_H - W_{H_2^+} = 2.62 \text{ v.e.}$$

No direct determination of this quantity has been made.

The value 2247 cm^{-1} for ω_0 for H_2^+ obtained by Birge (table 1) leads to $\frac{1}{2}\hbar\omega_0 = 0.141$ v.e., in very good agreement with the theoretical 0.14 v.e.

The average electron density ψ^2 as a function of position was also found by Burrau. For the normal state ψ^2 is represented in figure 6, in which the contour lines pass through points of relative density 1.0, 0.9, 0.8, . . . 0.1. The second curve represents the values of ψ^2 along the line passing through the two nuclei. (ψ^2 is cylindrically symmetrical about this line). It will be seen that the electron is most of the time in the region between the two nuclei, and can be considered as belonging to them both, and forming a bond between them.

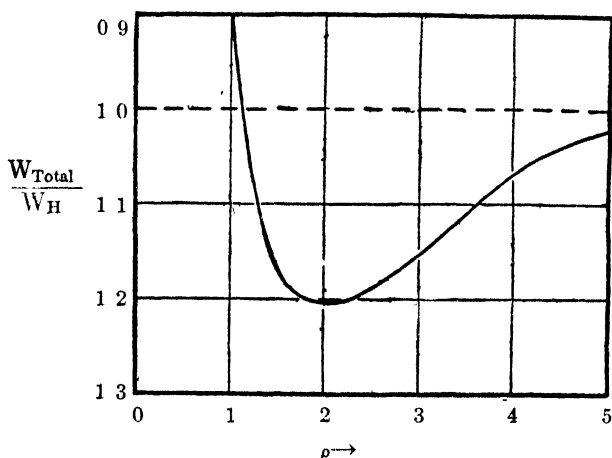


FIG. 5. THE TOTAL ENERGY OF THE HYDROGEN MOLECULE-ION AS A FUNCTION OF ρ (BURRAU)

b. Application of the first-order perturbation theory

Although no new numerical information regarding the hydrogen molecule-ion can be obtained by treating the wave equation by perturbation methods, nevertheless it is of value to do this. For perturbation methods can be applied to many systems for which the wave equation can not be accurately solved, and it is desirable to have some idea of the accuracy of the treatment. This can be gained from a comparison of the results of the perturbation method of the hydrogen molecule-ion and of Burrau's accurate numerical solution. The perturbation treatment assists, more-

over, in the physical interpretation of the forces operative in the formation of the ion; for the electronic energy is obtained as the sum of terms each of which can be related without difficulty to a visualisable process of interaction of the nuclei and the electron.

A hydrogen atom and a proton serve as the starting point of the calculation. With the nuclei a distance r_{AB} apart, and with the electron attached to the nucleus A to form a hydrogen atom, the zeroth order eigenfunction is

$$\psi = \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{\xi_A}{2}}, \quad \xi_A = \frac{2 Z_A}{a_0} r_A \quad (17a)$$

and the action of nucleus B on the electron is the perturbation.

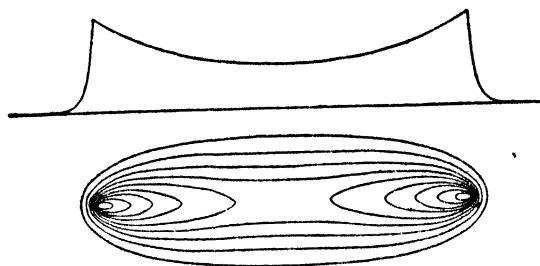


FIG. 6. THE ELECTRON DENSITY ψ^2 FOR THE HYDROGEN MOLECULE-ION IN THE NORMAL STATE (BURRAU)

The contour lines represent in section places of relative density 1.0, 0.9, 0.8, ..0.1. The density at points along a line drawn through the nuclei is given above.

But the configuration with the electron on nucleus B and with the eigenfunction

$$\varphi = \sqrt{\frac{Z^3}{\pi a_0^3}} e^{-\frac{\xi_B}{2}}, \quad \xi_B = \frac{2 Z_B}{a_0} r_B \quad (17b)$$

corresponds to the same energy. The unperturbed system is thus degenerate. There is, however, no perturbation function for the system as a whole, for in each case the interaction between the electron and the more distant nucleus produces the perturbation; and accordingly the usual theory for degenerate systems cannot

be directly applied. It is instead found that on carrying out the treatment used in the derivation of the ordinary perturbation theory closely similar results are obtained. The correct zeroth order eigenfunctions are linear aggregates of ψ and φ , and are, indeed, symmetric and antisymmetric in the two nuclei; they are

$$\Psi = \frac{1}{\sqrt{2 + 2S}} (\psi + \varphi) \quad (18a)$$

and

$$\Phi = \frac{1}{\sqrt{2 - 2S}} (\psi - \varphi) \quad (18b)$$

with

$$S = \int \psi \varphi \, d\Omega = e^{-\rho} (1 + \rho + \frac{1}{2} \rho^2), \quad (19a)$$

in which

$$\rho = \frac{Z}{a_0} r_{AB}, \text{ assuming } Z_A = Z_B = Z.$$

The radicals in the denominators are necessary in order that the new eigenfunctions be normalized. The wave equation (Equation 13) can now be written

$$\Delta\psi + \frac{8\pi^2 m}{h^2} \left(W_H + W^1 + \frac{Z_A e^2}{r_A} + \frac{Z_B e^2}{r_B} \right) \psi = 0, \quad (20)$$

in which W^1 is the perturbation energy. It is found (see Heitler and London (39) for the detailed treatment of a similar problem) that the perturbation energy for the eigenfunction Ψ is given by the solution of the equation

$$\int \Psi \left\{ \Delta \Psi + \frac{8\pi^2 m}{h^2} \left(W_H + W^1 + \frac{Z e^2}{r_A} + \frac{Z e^2}{r_B} \right) \Psi \right\} d\Omega = 0 \quad (21)$$

This leads to the result

$$W^1 = -e^2 \frac{I_1 + I_2}{1 + S} \quad (22)$$

in which

$$I_1 = \int \frac{\psi^2}{r_A} d\Omega = \frac{Z}{a_0} \left\{ \frac{1}{\rho} - e^{-2\rho} \left(1 + \frac{1}{\rho} \right) \right\} \quad (19b)$$

and

$$I_2 = \int \frac{\psi \varphi}{r_A} d\Omega = \frac{Z}{a_0} e^{-\rho} (1 + \rho). \quad (19c)$$

The total energy, including the internuclear term, is then

$$W_{\text{Total}} = W_{\text{H}} - e^2 \frac{I_1 + I_2}{1 + S} + \frac{e^2}{r_{AB}} \quad (23)$$

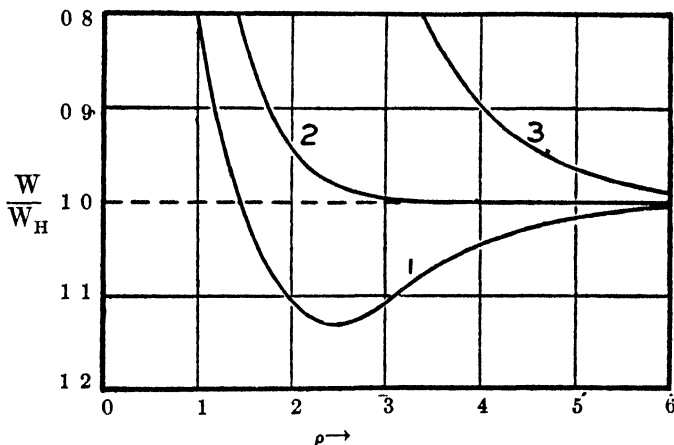


FIG. 7.

Curve 1 represents the total energy of the hydrogen molecule-ion as calculated by the first-order perturbation theory; curve 2, the naïve potential function obtained on neglecting the resonance phenomenon; curve 3, the potential function for the antisymmetric eigenfunction, leading to elastic collision.

In figure 7 is shown the so-calculated total energy W_{total} for H_2^+ as a function of ρ . Comparison with figure 5 shows that the perturbation curve is too high; the force holding the ion together is too small. Equilibrium occurs at $\rho = 2.5$ or $r_0 = 1.32 \text{ \AA}$ (correct value, 1.06 \AA), and the energy of the ion is then -15.30 v.e. (correct value, -16.30 v.e.).

The eigenfunction Ψ is represented qualitatively by figure 6, but is somewhat more extended, for the contracting influence of the second nucleus has not been taken into account.

It is of interest to calculate the energy corresponding to the single eigenfunction ψ (or φ); i.e., to find the interaction of a proton and a hydrogen atom that would result if the electron

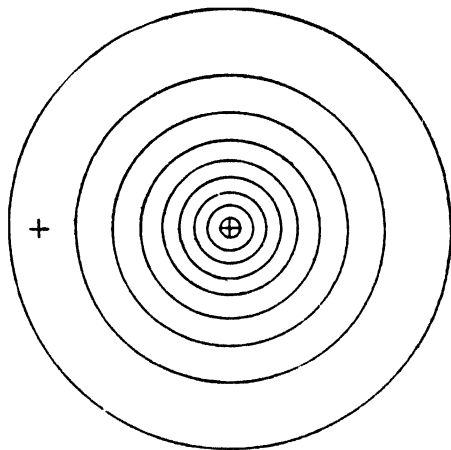


FIG. 8

The contour lines represent points of relative density 1.0, 0.9, 0.8, ... 0.1 for a hydrogen atom. This figure, with the added proton 1.06 Å from the atom, gives the electron distribution the hydrogen molecule-ion would have (in the zeroth approximation) if the resonance phenomenon did not occur, it is to be compared with figure 6 to show the effect of resonance.

were not allowed to jump from one nucleus to the other. For this case

$$W^1 = -e^2 I_1$$

and

$$W_{\text{Total}} = W_{\text{H}} - e^2 I_1 + \frac{e^2}{r_{AB}}. \quad (24)$$

(The perturbation energy is here just the electrostatic energy calculated for the electron distribution given by ψ^2 ; it is in part this feature of the perturbation theory which led Schrödinger to

proclaim that the electron is smeared through space, and which supports the present interpretation of ψ^2 as the probability that the electron will be in the selected region.) This naïve potential is also represented in figure 7. It leads to a repulsive force at all distances. The correct perturbation calculation given in Equation 23 differs from the naïve one in that it involves consideration of the interchange or resonance phenomenon, and leads to an attraction, giving an energy of formation of H_2^+ of about 1.62 v.e. (correct result from Burrau, 2.62 v.e.). We thus see that the fact that the electron can jump from one nucleus to the other—in other words, is shared between the two nuclei—is mainly responsible for the formation of the molecule-ion H_2^+ from H^+ and H . The way the electron distribution is affected by the resonance phenomenon is seen by comparing figure 6 with figure 8, which shows contour lines (of electron density ψ^2) for a hydrogen atom and a proton at the distance $\rho = 2$ (1.06 Å). This comparison is particularly effective in showing that the sharing of the electron between the two nuclei results from the resonance phenomenon.

This resonance energy leads to molecule formation only if the eigenfunction is symmetric in the two nuclei. The perturbation energy for the antisymmetric eigenfunction Φ is

$$W_A^1 = -e^2 \frac{I_1 - I_2}{1 - S}, \quad (25)$$

and the total energy,

$$W_{A \text{ Total}} = W_H - e^2 \frac{I_1 - I_2}{1 - S} + \frac{e^2}{r_{AB}}, \quad (26)$$

also shown in figure 7, leads to strong repulsion at all distances. This eigenfunction does not, then, give rise to a stable excited state of the hydrogen molecule-ion. Stable excited states will, however, correspond to the symmetric eigenfunctions relating to the various excited states of the hydrogen atom; and in each case (at least until deformation becomes very pronounced) not to the antisymmetric eigenfunctions. This is contrary to the qualitative scheme of levels given by Hund (31), who in a series of papers

(32) has considered the application of the quantum mechanics to molecules in general, with particular reference to molecular spectra.

The above perturbation treatment of the hydrogen molecule-ion has not before been published.

c. The second-order perturbation calculation

An attempt was made by Unsöld (33) to evaluate to the second-order the interaction of a proton and a hydrogen atom. He found, neglecting the resonance phenomenon, that the second-order perturbation energy is given approximately by the expression

$$W^2 = - \frac{9e^2}{4a_0} \left\{ \frac{1}{\rho^4} + e^{-2\rho} \left(-\frac{4}{3}\rho + 2 + \frac{4}{3\rho} + \frac{1}{\rho^2} + \frac{2}{\rho^3} + \frac{1}{\rho^4} \right) - e^{-4\rho} \left(1 + \frac{2}{\rho} + \frac{1}{\rho^2} \right) + \frac{8}{3}\rho^2 E_i(-2\rho) \right\}, \quad (27)$$

in which $E_i(-2\rho)$ represents the integral logarithm with the argument -2ρ . This expression is accurate for large values of ρ ; for it then reduces to $-\frac{1}{2} \frac{e^2\alpha}{r_{AB}^4}$, in which α , the polarizability of a hydrogen atom, has the value 0.667×10^{-24} deduced from the second-order Stark effect energy (34).

Using this expression and the value given in Equation 24 for the first-order perturbation, Unsöld found that equilibrium would occur at $\rho = 1$, $r_0 = 0.53 \text{ \AA}$; and that the electronic energy of the hydrogen molecule-ion would then be $-1.205 W_H$, or -16.31 v.e. , in exact agreement with Burrau's value. This agreement is, however, misleading, and indeed the calculation is of no significance, for Unsöld neglected to consider the resonance phenomenon, making use instead of the naïve first-order perturbation. We may, however, attempt to rectify this by adding the second-order energy W^2 to the correct first-order energy of Equation 23. When this is done it is found that equilibrium occurs at $\rho = 1.2$, $r_0 = 0.64 \text{ \AA}$, and that the energy then is -17.95 v.e.

The results of Burrau's calculation, of the first-order perturbation treatment, and of the second-order treatment are given in table 3.

It will be seen that the second-order treatment leads to results which deviate more from the correct values than do those given by the first-order treatment alone. This is due in part to the fact that the second-order energy was derived without consideration of the resonance phenomenon, and is probably in error for that reason. The third-order energy is also no doubt appreciable. It can be concluded from table 3 that the first-order perturbation calculation in problems of this type will usually lead to rather good results, and that in general the second-order term need not be evaluated.

TABLE 3
The properties of the hydrogen molecule-ion

	ρ	r_0	$W_{H_2^+}$
Burrau.	2 0	1.06 Å	-16 30 v.e.
First-order.	2 5	1 32	-15.30
Second-order	1 2	0.64	-17 95

In dealing with problems of molecular structure it has been customary in the past to introduce the energy of polarization of one atom or ion by another ion in the form $-\frac{1}{2} \frac{e^2 \alpha}{r_{AB}^4}$, and to give α the value it possesses in a uniform field. The form of Equation 27 shows that this is not a good representation of the polarization energy, for it gives values which are much too large at small distances. Indeed, if we attach this term to the first-order energy of Equation 23, it is found that the correct equilibrium distance $r_0 = 1.06 \text{ Å}$ results only if α be placed equal to 0.032×10^{-24} , which is only 5 per cent of the true polarizability of hydrogen. This indicates that in general better results will be obtained in the theoretical treatment of the structure of molecules by ignoring polarization completely than by introducing the usual expression for the polarization. As a matter of fact, it has already been pointed out by Fajans (35) that the experimental values of the heat of sublimation of the alkali halides agree better with

those calculated by Reis (36), who neglected polarization, than with those of Born and Heisenberg (37), who took it into consideration. As a result of this it is probable that the numerous discussions of molecular structure based on polarization (38) are of only limited significance.

VII. THE HYDROGEN MOLECULE

a. Perturbation treatment of the interaction of two hydrogen atoms

The most satisfactory treatment which has been given the structure of the hydrogen molecule is that of Heitler and London



FIG. 9. DIAGRAMMATIC REPRESENTATION OF THE EIGENFUNCTIONS FOR TWO HYDROGEN ATOMS

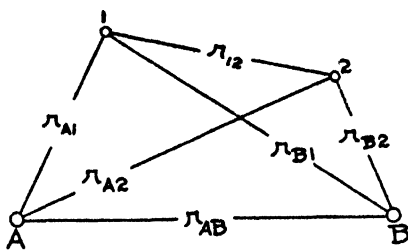


FIG. 10. DIAGRAM SHOWING THE SYMBOLS USED FOR THE VARIOUS INTERNUCLEAR AND INTERELECTRONIC DISTANCES

(39). The system to be treated consists of two nuclei A and B and two electrons 1 and 2. In the unperturbed state two hydrogen atoms are assumed, so that the zeroth-order energy is $2W_H$. If the first electron is attached to nucleus A and the second to nucleus B , the zeroth-order eigenfunction is $\psi(1)\varphi(2)$, in which ψ and φ are the functions given in Equations 17a and 17b. The state obtained by interchanging the two electrons, $\varphi(1)\psi(2)$, corresponds to the same zeroth-order energy, so that the system

is degenerate, and must be treated in a way similar to that applied to the hydrogen molecule-ion in Section VIb.

The wave equation for the hydrogen molecule with fixed nuclei is

$$\Delta_1 \psi + \Delta_2 \psi + \frac{8\pi^2 m}{h^2} \left\{ 2 W_H + W^1 + \frac{e^2}{r_{A1}} + \frac{e^2}{r_{B1}} + \frac{e^2}{r_{A2}} + \frac{e^2}{r_{B2}} - \frac{e^2}{r_{12}} - \frac{e^2}{r_{AB}} \right\} \psi = 0 \quad (28)$$

in which Δ_1 and Δ_2 represent the second differential operations with respect to the coördinates of the first and the second electron respectively, and the various r 's refer to the distances indicated in figure 10. The correct zeroth-order eigenfunctions are found by the procedure of Section VIb to be the symmetric

$$\Psi_{H_2} = \frac{1}{\sqrt{2 + 2S^2}} \{ \psi(1) \varphi(2) + \varphi(1) \psi(2) \} \quad (29a)$$

and the antisymmetric

$$\Phi_{H_2} = \frac{1}{\sqrt{2 - 2S^2}} \{ \psi(1) \varphi(2) - \varphi(1) \psi(2) \} \quad (29b)$$

in which S is given by Equation 19a.

In this case, too, molecule formation results from the symmetric eigenfunction. The corresponding perturbation energy W^1 is obtained from an equation of the type of Equation 20 involving Ψ_{H_2} and the wave equation 28. It is

$$W^1 = \frac{e^2}{1 + S^2} \left\{ \frac{1 + S^2}{r_{AB}} + I_4 + I_6 - 2I_1 - 2SI_2 \right\} \quad (30)$$

in which I_1 and I_2 are given by Equations 19b and 19c, and I_4 and I_6 by

$$I_4 = \iint \frac{\psi^2(1) \varphi^2(2)}{r_{12}} d\Omega_1 d\Omega_2 = \frac{Z}{a_0} \left\{ \frac{1}{\rho} - e^{-2\rho} \left(\frac{1}{\rho} + \frac{11}{8} + \frac{3}{4}\rho + \frac{1}{6}\rho^2 \right) \right\} \quad (19d)$$

and

$$I_6 = \iint \frac{\psi(1) \varphi(1) \psi(2) \varphi(2)}{r_{12}} d\Omega_1 d\Omega_2$$

$$= \frac{Z}{5a_0} \left[-e^{-2\rho} \left(-\frac{25}{8} + \frac{23}{4}\rho + 3\rho^2 + \frac{1}{3}\rho^3 \right) \right. \quad (31)$$

$$\left. + \frac{6}{\rho} \left\{ S^2 (\gamma + \log \rho) + S'^2 E_1(-4\rho) - 2SS' E_1(-2\rho) \right\} \right], \quad (19e)$$

in which $\gamma = 0.5772 \dots$ is Euler's constant, and $S' = e^\rho (1 - \rho + \frac{1}{3}\rho^2)$.

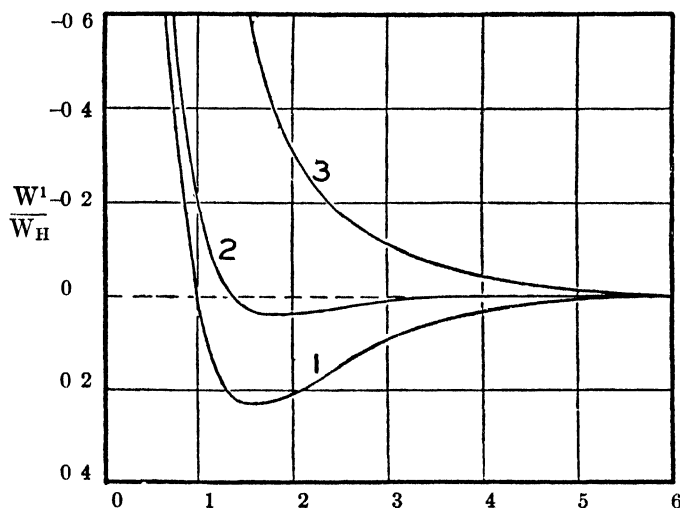


FIG. 11

Curve 1 shows the total energy for the normal state of the hydrogen molecule as given by the first-order perturbation theory; curve 2, the naive potential function obtained by neglecting the resonance phenomenon; and curve 3, the potential function for the antisymmetric eigenfunction, corresponding to elastic collision.

The energy W^1 depends largely on the integral I_6 , for which Heitler and London gave only an approximation. The difficult problem of carrying out this integration was solved by Sugiyra (40), whose result is given in Equation 19e. W^1 is shown as a function of ρ in figure 11. It has a minimum at the equilibrium distance $\rho = 1.5$, at which $W^1 = -3.1$ v.e. The comparison of theory and experiment for the hydrogen molecule is shown in

table 4. The general agreement is entirely satisfactory in view of the fact that only the first-order perturbation calculation has been made.

It is worthy of mention that the perturbation terms are actually larger than appears from W^1 . Thus at $\rho = 1.5$ the term $\frac{e^2}{r_{AB}}$ has the value 18.1 v.e., so that the electronic perturbation energy is -21.3 v.e., which is to be compared with the experimental value -22.5 v.e. It is seen that the electronic perturbation energy is thus in error by only 5 per cent. Furthermore, we saw in Section VIb that the first-order perturbation calculation for the hydrogen molecule-ion gave an energy of dissociation of H_2^+ of 1.62 v.e., which is 1 v.e. smaller than Burrau's correct value. A similar error is to be expected for the hydrogen molecule; and, as a matter of fact, the calculated energy of dissociation is here 1.2 v.e. too small. We are hence justified in the belief that the

TABLE 4
Properties of the hydrogen molecule

	r_0	I_0	D_{H_2}	ω_0
Calculated.....	0 80 Å	0.53×10^{-40} g. cm. ²	3.2 v.e.	4800 cm. ⁻¹
Observed	0.76	0.48	4.42	4262

accurate theoretical treatment of the hydrogen molecule would give results in complete agreement with experiment.

By bringing the nuclei into coincidence a helium atom in the normal state is formed; and a value for its energy can be obtained from the expression for the hydrogen molecule by neglecting the internuclear energy and by putting $\rho = 0$. It is found that W_{He}

$= -\frac{19}{8}e^2/a_0 = -64.3$ v.e., which lies considerably above the experimental value -78.8 v.e.; the error is in the same direction as that for $D_{H_2^+}$. A similar limiting calculation for the hydrogen molecule-ion gives $W_{He^+} = -3e^2/2a_0 = -40.6$ v.e., instead of the correct value $-4W_H = -54.16$ v.e. Thus for both He and He^+ this perturbation treatment is inaccurate. The treatment

gives, however, $64.3 - 40.6 = 23.7$ v. for the first ionisation potential of helium, which is in good agreement with the experimental 24.6 v. This agreement was adduced by Suguira as evidence of the accuracy of the perturbation treatment; but it is merely accidental.

If $\psi(1)\varphi(2)$ or $\varphi(1)\psi(2)$ alone were to be considered, the perturbation energy would be the electrostatic energy of two nuclei and two electrons distributed according to the probability functions ψ^2 and φ^2 ; namely,

$$W^1 = e^2 \left\{ \frac{1}{r_{AB}} + I_1 - 2I_2 \right\} = \frac{e^2}{a_0} e^{-2\rho} \left(\frac{1}{\rho} + \frac{5}{8} - \frac{3}{4}\rho - \frac{1}{6}\rho^2 \right) \quad (31)$$

This, the naïve potential function, is also shown in figure 11. It corresponds to a relatively small attraction, so that the conclusion can be drawn that in the hydrogen molecule the interchange energy of the two electrons is the principal cause of the forces leading to molecule formation.

b. The application to the hydrogen molecule of Burrau's numerical solution for the molecule-ion

If the interaction of the two electrons in the hydrogen molecule were small it could be neglected with respect to the electronic-nuclear interaction. Each electron would then be represented by an H_2^+ eigenfunction, and the electronic energy would be just twice that calculated by Burrau. The interelectronic energy could then be calculated as a perturbation; this would necessitate, however, a knowledge of Burrau's eigenfunctions in a form suitable for integration.

Lacking this knowledge, Condon (41) made use of the following semi-empirical method. The electronic energy of the helium atom is raised from $-4 e^2/a_0 = -108.32$ v.e. (Burrau's value) to -78.8 v.e. (the experimental value) by the electronic interaction. If it be assumed that the interelectronic energy is to be raised to the same fractional value of the Burrau energy throughout the region corresponding to the normal state of the hydrogen molecule, there is obtained the electronic energy function shown in figure 12. Adding to this the electronic energy e^2/r_{AB} , it is

found that equilibrium occurs at $r_0 = 0.73 \text{ \AA}$, with $D_{\text{H}_2} = 4.3$ or 4.4 v.e. and $\omega_0 = 7600 \text{ cm.}^{-1}$. These values, except for ω_0 , are in excellent agreement with the experimental ones.

This treatment is, however, of less significance than that of the preceding section, for it is more arbitrary and less confidence can be placed in it. In the first place, the interelectronic forces are not small, and it is to be expected that they will cause the electrons to tend to remain on different atoms, as is assumed in the previous

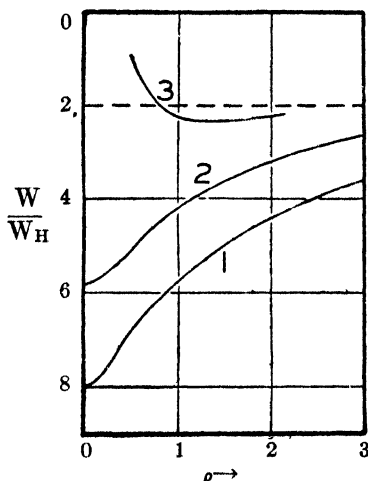


FIG. 12

Curve 1 shows the electronic energy of the hydrogen molecule neglecting inter-electronic interaction (from Burrau's solution for the molecule-ion); curve 2, the electronic energy empirically corrected by Condon's method; and curve 3, the total energy of the hydrogen molecule, calculated by Condon's method.

treatment of Heitler and London. The assumption that the total electronic energy as calculated by Burrau should be reduced in a constant ratio is, moreover, without justification. It is definitely incorrect for ρ large; for the doubled Burrau energy then approaches the correct value $2W_{\text{H}}$. An assumption which might just as well be made and which is satisfactory both for $\rho = 0$ and for $\rho = \infty$ is that the difference between the doubled Burrau energy and $2W_{\text{H}}$ is to be reduced in a constant ratio; but it leads to the incorrect values $D_{\text{H}_2} = 8.0 \text{ v.e.}$ and $r_0 = 0.90 \text{ \AA}$.

c. The excited states of the hydrogen molecule

Excited states of the hydrogen molecule may be formed from a normal hydrogen atom and a hydrogen atom in various excited states.² For these the interelectronic interaction will be small, and the Burrau eigenfunction will represent the molecule in part with considerable accuracy. The properties of the molecule, in particular the equilibrium distance, should then approximate those of the molecule-ion; for the molecule will be essentially a molecule-ion with an added electron in an outer orbit. This is observed in general; the equilibrium distances for all known excited states but one (the second state in table 1) deviate by less than 10 per cent from that for the molecule-ion. It is hence probable that states 3, 4, 5, and 6 are formed from a normal and an excited atom with $n = 2$, and that higher states are similarly formed.

The exceptional state *B* has a very large equilibrium distance and small oscillational frequency, as has been pointed out by Birge (7). This suggests that the molecule is here not non-polar, but is a polar compound of H^+ and H^- . The electron affinity of hydrogen is probably negative, (about -1 kcal/mole (42)), and it is doubtful that a free negative hydrogen ion in the normal state can exist. The presence of another proton would, however, stabilize the structure, so that a polar hydrogen molecule could be formed. The unperturbed system is again degenerate, for both electrons may be attached to nucleus A or to nucleus B. The zeroth-order eigenfunction representing the most stable polar state of the molecule is

$$\frac{1}{\sqrt{2 + 2 S^2}} \{ \varphi(1) \varphi(2) + \psi(1) \psi(2) \}. \quad (32)$$

The first-order perturbation theory in this case does not give good results, since the mutual interaction of two electrons on one nucleus is so large as to greatly deform the eigenfunctions; it leads

² The calculation of the potential function for these states with the use of the method of Heitler and London is being made by Prof. E. C. Kemble (private communication to Dr. J. R. Oppenheimer).

to an equilibrium distance of about 1\AA and an energy about 5 v.e. greater than the normal state. There is no energy level of the molecule in this region. The attempt to take into account the effect of deformation has led to the conclusion that both r_0 and the energy should be increased to values compatible with those observed for the first excited state B . Since a polar state is to be expected in this region and since B has properties explicable on

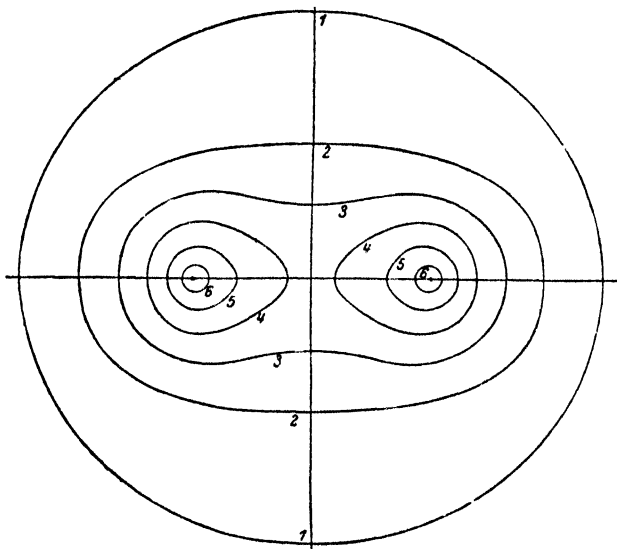


FIG. 13 QUALITATIVE REPRESENTATION BY CONTOURS OF ELECTRON DISTRIBUTION FOR TWO HYDROGEN ATOMS UNITING TO FORM A MOLECULE (LONDON)

this basis but not otherwise, the identification of the two may be made with some certainty.

We shall next consider whether or not the antisymmetric eigenfunction Φ_H , for two hydrogen atoms (Equation 29b) would lead to an excited state of the hydrogen molecule. The perturbation energy is found to be

$$W^1 = \frac{e^2}{1 - S^2} \left\{ \frac{1 - S^2}{r_{AB}} + I_4 - I_6 - 2 I_1 + 2 S I_2 \right\} \quad (33)$$

This potential differs from that of Equation 30 in that the **interchange** energy has the opposite sign (and slightly different **magnitude**). As a result it corresponds to repulsion between the two atoms at all distances, and not to a stable state of the molecule (see figure 11). This result, which had been suggested as a possibility by Hund (31), was proved by Heitler and London. The existence of two potential functions representing the interaction of two normal hydrogen atoms is very remarkable and has, I believe, no classical interpretation. A certain feeling

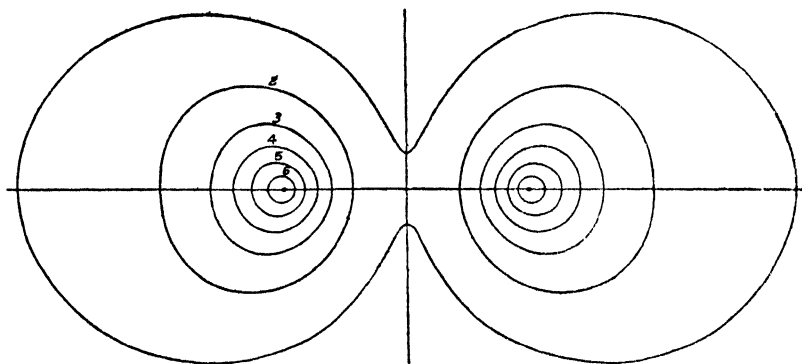


FIG 14. THE ELECTRON DISTRIBUTION FOR TWO HYDROGEN ATOMS IN ELASTIC COLLISION (LONDON)

for the phenomenon results from the study of the distribution of the two electrons in the two cases (London, 46). In figure 13, representing the two hydrogen atoms in the process of forming a molecule, it is seen that the electrons tend to assume positions between the two nuclei, and form a bond between them—the shared electron-pair. But if the potential function represents elastic collision (the antisymmetric eigenfunction) the electrons take up positions on the outer sides of the atoms (fig. 14), with the result that the strong internuclear repulsion becomes effective.

VIII. THE PAULI EXCLUSION PRINCIPLE. THE INTERACTION OF TWO HELIUM ATOMS

It was mentioned in Section III that the fine-structure of spectra arises from a phenomenon equivalent to a first approxi-

mation to that resulting from a spin of the electron. The spin moment of the electron can assume two orientations in space, which may be represented by spin eigenfunctions α and β . In the foregoing sections discussion has been given only to eigenfunctions referring to the electronic positions. A complete eigenfunction is the product of a positional eigenfunction and a spin eigenfunction, as $\psi \alpha$ or $\psi \beta$.

The observed structure of the spectra of many-electron atoms is entirely accounted for by the following postulate: *Only eigenfunctions which are antisymmetric in the electrons; that is, change sign when any two electrons are interchanged, correspond to existant states of the system.* This is the quantum mechanics statement (26) of the Pauli exclusion principle (43).

It is equivalent to saying that two electrons cannot occupy the same orbit. Thus there is no antisymmetric eigenfunction composed of $\psi(1) \alpha(1)$ and $\psi(2) \alpha(2)$, and no such state exists (for the helium atom, say). The allowed state is $\frac{1}{\sqrt{2}} \{ \psi(1) \alpha(1) \psi(2) \beta(2) - \psi(1) \beta(1) \psi(2) \alpha(2) \}$; that is, in the normal state of the helium atom the two electrons have oppositely directed spins. Other consequences of the exclusion principle, such as that not more than two electrons can occupy the K-shell of an atom, follow directly.

In dealing with systems containing only two electrons we have not been troubled with the exclusion principle, but have accepted both symmetric and antisymmetric positional eigenfunctions; for by multiplying by a spin eigenfunction of the proper symmetry character an antisymmetric total eigenfunction can always be obtained. In the case of two hydrogen atoms there are three symmetric spin eigenfunctions $\alpha(1) \alpha(2)$, $\beta(1) \beta(2)$, and $\frac{1}{\sqrt{2}} \{ \alpha(1) \beta(2) + \beta(1) \alpha(2) \}$, and one antisymmetric, $\frac{1}{\sqrt{2}} \{ \alpha(1) \beta(2) - \beta(1) \alpha(2) \}$. The last is required to make the symmetric positional eigenfunction Ψ_{H_2} of Equation 29a conform to Pauli's principle, and the first three for the antisymmetric Φ_{H_2} . Since the a priori probability of each eigenfunction is the same, there

are three chances that two hydrogen atoms will repel each other to one that they will attract.

But if the system contains more than two electrons explicit consideration must be given the spins. This is particularly evident in the problem of the interaction of two helium atoms. There are four individual eigenfunctions $\psi\alpha$, $\psi\beta$, $\varphi\alpha$, and $\varphi\beta$, which are to be occupied by the four electrons. The only eigenfunction allowed by Pauli's principle for the system is

$$\Psi_{\text{He}_2} = a \begin{vmatrix} \psi(1) \alpha(1) & \psi(1) \beta(1) & \varphi(1) \alpha(1) & \varphi(1) \beta(1) \\ \psi(2) \alpha(2) & \psi(2) \beta(2) & \varphi(2) \alpha(2) & \varphi(2) \beta(2) \\ \psi(3) \alpha(3) & \psi(3) \beta(3) & \varphi(3) \alpha(3) & \varphi(3) \beta(3) \\ \psi(4) \alpha(4) & \psi(4) \beta(4) & \varphi(4) \alpha(4) & \varphi(4) \beta(4) \end{vmatrix} \quad (34)$$

(a is a factor of such value as to make the eigenfunction normalized.)

It will be seen that this is antisymmetric, for interchanging any two electrons is equivalent to interchanging two rows of the determinant, and hence to changing its sign.

Substitution of this eigenfunction in an expression of the type of Equation 21 permits the evaluation of the perturbation energy W^1 , in the course of which use is made of the properties of orthogonality and normalization of the spin eigenfunctions; namely,

$$\left. \begin{aligned} \int \alpha^2 d\epsilon &= \int \beta^2 d\epsilon = 1 \\ \int \alpha \beta d\epsilon &= 0 \end{aligned} \right\} \quad (35)$$

in which ϵ is the variable occurring in the spin eigenfunctions. The value of one further integral is also needed:

$$\begin{aligned} I_1 &= \iint \frac{\psi^2(1) \psi(2) \varphi(2)}{r_{12}} d\Omega_1 d\Omega_2 \\ &= \frac{Z}{a_0} \left\{ e^{-\rho} \left(\frac{5}{16\rho} + \frac{1}{8} + \rho \right) - e^{-3\rho} \left(\frac{5}{16\rho} + \frac{1}{8} \right) \right\}. \end{aligned} \quad (19f)$$

The potential function obtained is only approximately correct, for the eigenfunctions are in fact largely perturbed by the inter-electronic interaction. There are no forces tending to molecule

formation, but instead repulsion at all distances. The van der Waals' attractive force (which is very small for helium) does not appear, and the repulsive force is much larger than the actual one; these discrepancies are no doubt due to the inaccuracy of the calculation.

It is of interest to carry out the evaluation of the naïve potential function obtained from the eigenfunction $\psi(1)\alpha(1)\psi(2)\beta(2)\varphi(3)\alpha(3)\varphi(4)\beta(4)$; i.e., with the neglect of the interchange energy of the electrons. This potential leads to a strong attractive force, with the formation of molecules He_2 with about 10,000 or 15,000 cal/mole dissociation energy. The resonance phenomenon is accordingly largely responsible for the very small van der Waals' forces in helium; without it the boiling point of helium would be around room temperature.

IX. OTHER RELATED PROBLEMS. THE EXTENSION OF THE THEORY

The interaction of two alkali metal atoms is to be expected to be similar to that of two hydrogen atoms, for the completed shells of the ions will produce forces similar to the van der Waals' forces of a rare gas. The two valence electrons, combined symmetrically, will then be shared between the two ions, the resonance phenomenon producing a molecule-forming attractive force. This is, in fact, observed in band spectra. The normal state of the Na_2 molecule, for example, has an energy of dissociation of 1 v.e. (44). The first two excited states are similar, as is to be expected; they have dissociation energies of 1.25 and 0.6 v.e. respectively.

In an atom of the second column of the periodic system, such as mercury, the two valence electrons are in the normal state s-electrons, and form a completed sub-group. Two such atoms would hence interact in a way similar to two helium atoms; the attractive forces would be at most very small. This is the case for Hg_2 , which in the normal state has an energy of dissociation of only 0.05 v.e. But if one or both of the atoms is excited strong attractive forces can arise; and indeed the excited states of Hg_2 are found to have energies of dissociation of about 1 v.e.

Similarly two Hg^+ ions will attract each other with some force to form the stable Hg_2^{++} ion long recognized by chemists.

The application of the quantum mechanics to the interaction of more complicated atoms, and to the non-polar chemical bond in general, is now being made (45). A discussion of this work can not be given here; it is, however, worthy of mention that qualitative conclusions have been drawn which are completely equivalent to G. N. Lewis's theory of the shared electron pair. The further results which have so far been obtained are promising; and we may look forward with some confidence to the future explanation of chemical valence in general in terms of the Pauli exclusion principle and the Heisenberg-Dirac resonance phenomenon.

NOTE ADDED IN PROOF

Since the submission of this article for publication a number of pertinent papers have appeared.

H. A. Wilson (*Proc. Roy. Soc. London*, A **118**, 635 (1928); see also (29)) states that no functions satisfying the wave equation for the hydrogen molecule-ion and bounded everywhere exist. There are, however, solutions which become logarithmically infinite along the nuclear axis and are bounded elsewhere. These solutions would not be considered eigenfunctions if the usual definition is retained; but would be in case the restriction that the eigenfunction be bounded everywhere were replaced by the restriction that it be quadratically integrable. Wilson has made this assumption, and has found that the so-calculated properties of the hydrogen molecule-ion in the normal state are approximately those given by Burrau. An accurate treatment and the consideration of excited states have not been published.

A treatment of the hydrogen molecule by the Ritz method, applied to helium by Kellner (25), has been reported by S. C. Wang (*Phys. Rev.*, **31**, 579 (1928)). With this method the individual eigenfunctions ψ and φ (equation 29) are taken to be the hydrogen-like eigenfunctions of an atom with atomic number Z differing from unity. The value found for Z is 1.166, and the

corresponding constants of the hydrogen molecule in the normal state are

$$r_0 = 0.75 \text{ \AA}, I_0 = 0.459 \times 10^{-40} \text{ g. cm.}^2, D_{H_2} = 3.76 \text{ v.e.}, \omega_0 = 4900 \text{ cm}^{-1}.$$

Comparison with table 4 shows that these values are in somewhat better agreement with the observed ones than are Sugiura's.

B. N. Finkelstein and G. E. Horowitz (*Z. f. Physik*, **48**, 118 (1928)) have similarly applied the Ritz method to the hydrogen molecule-ion, obtaining the following values:

$$Z = 1.228, \rho \cong 2, r_0 \cong 1.06 \text{ \AA}, W_{H_2^+} = -15.75 \text{ v.e.}$$

These results are better than those given by the perturbation theory (table 3).

F. R. Bichowsky and I. C. Copeland (*J. Am. Chem. Soc.*, **50**, 1315 (1928)) have made a direct determination of the heat of formation of molecular hydrogen, leading to the value $D_{H_2} = 105000 \pm 3500 \text{ cal/mole}$.

REFERENCES

- (1) J. H. VAN VLECK, "Quantum Principles and Line Spectra," *Bull. of the Nat. Res. Council*, No. 54, **1926**, Chap. VII.
See also E. HUTCHISSON, *Phys. Rev.*, **29**, 270 (1927); J. F. NIESSEN, *Z. f. Physik*, **43**, 694 (1927).
- (2) W. HEISENBERG, *Z. f. Physik*, **33**, 879 (1925).
- (3) M. BORN AND P. JORDAN, *Z. f. Physik*, **34**, 858 (1925).
M. BORN, W. HEISENBERG, AND P. JORDAN, *Ibid.*, **35**, 557 (1926); etc.
- (4) E. SCHRÖDINGER, I. *Ann. d. Physik*, **79**, 361 (1926); II, *Ibid.*, **79**, 489 (1926); III, *Ibid.*, **80**, 437 (1926); IV, *Ibid.*, **81**, 109 (1926); *Phys. Rev.*, **28**, 1049 (1926).
- (5) E. SCHRÖDINGER, *Ann. d. Physik*, **79**, 734 (1926).
C. ECKART, *Phys. Rev.*, **28**, 711 (1926).
W. PAULI, *Unpublished work*.
- (6) P. A. M. DIRAC, *Proc. Roy. Soc. A.*, **110**, 561 (1926); *Ibid.*, **111**, 281 (1926); *Ibid.*, **112**, 661 (1926); *Ibid.*, **113**, 621 (1927); *Ibid.*, **114**, 243, 610 (1927).
P. JORDAN, *Z. f. Physik*, **40**, 809 (1927); **44**, 1 (1927).
W. HEISENBERG, *Z. f. Physik*, **43**, 172 (1927).
- (7) R. T. BIRGE, *Proc. Nat. Acad.*, **14**, 12 (1928); *Nature*, **121**, 134 (1928).
- (8) G. H. DIEKE AND J. J. HOPFIELD, *Z. f. Physik*, **40**, 299 (1926), *Phys. Rev.*, **30**, 400 (1927).
- (9) E. E. WITMER, *Phys. Rev.*, **28**, 1223 (1926).
- (10) O. W. RICHARDSON, *Proc. Roy. Soc., London, A.*, **111**, 714 (1926); **113**, 368 (1926); **114**, 643 (1927); **115**, 528 (1927); *Nature*, **121**, 320 (1928).

- (11) J. FRANCK AND P. JORDAN, "Die Anregung von Quantensprünge durch Stosse," Berlin, 1926
- (12) J. FRANCK, *Trans Faraday Soc*, **21**, 536 (1925).
E. U CONDON, *Phys Rev*, **28**, 1182 (1926); *Proc. Nat. Acad*, **13**, 462 (1927).
- (13) T. ISNARDI, *Z. f. Elektrochemie*, **21**, 405 (1915).
- (14) K WOHL, *Z. f. Elektrochemie*, **30**, 49 (1924).
- (15) I. LANGMUIR, *Gen. Elec. Rev.*, **29**, 153 (1926).
- (16) T. HORI, *Z. f. Physik*, **44**, 834 (1927).
- (17) D. M. DENNISON, *Proc. Roy. Soc. London, A.*, **115**, 483 (1927).
- (18) R. E. CORNISH AND E. D. EASTMAN, *J. Am. Chem. Soc*, **50**, 627 (1928).
- (19) I. WALLER, *Z. f. Physik*, **38**, 635 (1926).
- (20) See, for example, COURANT-HILBERT, "Methoden der Mathematischen Physik," Berlin, 1924
- (21) See WHITTAKER AND WATSON, "Modern Analysis."
BYERLY, "Fourier's Series and Spherical Harmonics," etc.
- (22) P. A. M. DIRAC, *Proc. Roy. Soc. London, A.*, **117**, 610 (1928).
- (23) G. E. UHLENBECK AND S. GOUDSMIT, *Naturwissenschaften*, **13**, 953 (1925).
Nature, **107**, 264 (1926).
- (24) A. UNSOLD, *Ann. d. Physik*, **82**, 355 (1927).
- (25) G. W. KELLNER, *Z. f. Physik*, **44**, 91 (1927); **44**, 110 (1927).
- (26) W. HEISENBERG, *Z. f. Physik*, **38**, 411 (1926); **39**, 499 (1926); **41**, 239 (1927).
P. A. M. DIRAC, *Proc. Roy. Soc. London, A*, **112**, 661 (1926).
- (27) M. BORN AND J. R. OPPENHEIMER, *Ann. d. Physik*, **84**, 457 (1927).
- (28) W. ALEXANDROW, *Ann. d. Physik*, **81**, 603 (1926).
- (29) A. H. WILSON, *Nature*, **121**, 304 (1928).
- (30) O. BURRAU, Det. Kgl. Danske Videnskabernes Selskab Math.-fys. Meddelelser, VII, 14 (1927).
- (31) F. HUND, *Z. f. Physik*, **40**, 742 (1927).
- (32) F. HUND, *Z. f. Physik*, **42**, 93 (1927); **43**, 805 (1927).
- (33) A. UNSOLD, *Z. f. Physik*, **43**, 563 (1927)
- (34) G. WENTZEL, *Z. f. Physik*, **38**, 518 (1926).
I. WALLER, *Ibid*, **38**, 635 (1926).
P. EPSTEIN, *Phys. Rev.*, **28**, 695 (1926).
J. H. VAN VLECK, *Proc. Nat. Acad*, **12**, 662 (1926).
L. PAULING, *Proc. Roy. Soc. London, A*, **114**, 181 (1927).
- (35) K. FAJANS, *Private communication to the author.*
- (36) A. REIS, *Z. f. Physik*, **1**, 294 (1920).
- (37) M. BORN AND W. HEISENBERG, *Z. f. Physik*, **23**, 388 (1924).
- (38) W. HEISENBERG, *Z. f. Physik*, **26**, 196 (1924).
H. KORNFIELD, *Ibid*, **26**, 205 (1924).
F. HUND, *Ibid.*, **31**, 81 (1925); **32**, 1 (1925).
V. GUILLEMIN, *Ann. d. Physik*, **81**, 173 (1926).
J. H. DE BOER AND A. E. VAN ARKEL, *Z. f. Physik*, **41**, 27, 38 (1927); etc
- (39) W. HEITLER AND F. LONDON, *Z. f. Physik*, **44**, 455 (1927).
- (40) Y. SUGIURA, *Z. f. Physik*, **45**, 484 (1927).
- (41) E. U. CONDON, *Proc. Nat. Acad*, **13**, 466 (1927).
- (42) L. PAULING, *Phys. Rev*, **29**, 285 (1927).
J. KASARNOWSKY AND M. PROSKURNIN, *Z. f. Physik*, **43**, 512 (1927).

- (43) W. PAULI, *Z. f. Physik*, **31**, 765 (1925).
- (44) F. W. LOOMIS, *Phys. Rev*, **31**, 323 (1928).
- (45) W. HEITLER, *Z. f. Physik*, **46**, 47 (1927); *Ibid.*, **47**, 835 (1928).
Nachrichten d. Ges. d. Wiss. z. Göttingen, **1927**, p. 368.
F. LONDON, *Z. f. Physik*, **46**, 455 (1928).
L. PAULING, *Proc. Nat. Acad.*, **14**, 359 (1928).

THE HELIX CHEMICA

B. K. EMERSON

Amherst, Massachusetts

In 1911 I devised a helical arrangement of the elements (1) which, being a three-dimensional projection as compared with the two-dimensional table of Mendeleeff, still seems the best for presenting the multiform relations of the periodic system and for predicting new relationships. The helix has an exceptional value as a scheme of representation since it can be drawn so simply, can be viewed from the side as well as from the end, and because it expresses a fundamental harmony in a spiral whose coils increase in size and number according to the simplest geometrical ratio of the powers of two while the interspaces (representing interweights) increase by the simplest arithmetical series, 1, 2, 3, 4, etc.

The purpose of this paper is to show that the addition of two small coils at the origin of the helix as originally published perfects it and that the full development of the series of the powers of two demands this addition. It requires one additional element before hydrogen but this addition completes the symmetry of the elemental series. The same symmetry suggests the existence of two new elements between hydrogen and helium as well as the two preceding hydrogen.

THE ORIGINAL HELIX

In constructing the helix in 1911, as shown in figure 1, I started with the recognition of two symmetrical groups, of eight elements each, in two circles,—the octaves,—followed by four similar groups of sixteen elements each (counting the triads as one) in four circles,—the double octaves,—and the first quadrant of a larger circle with 32 elements,—the quadruple octave. This symmetry demanded at the other end, going backward from the

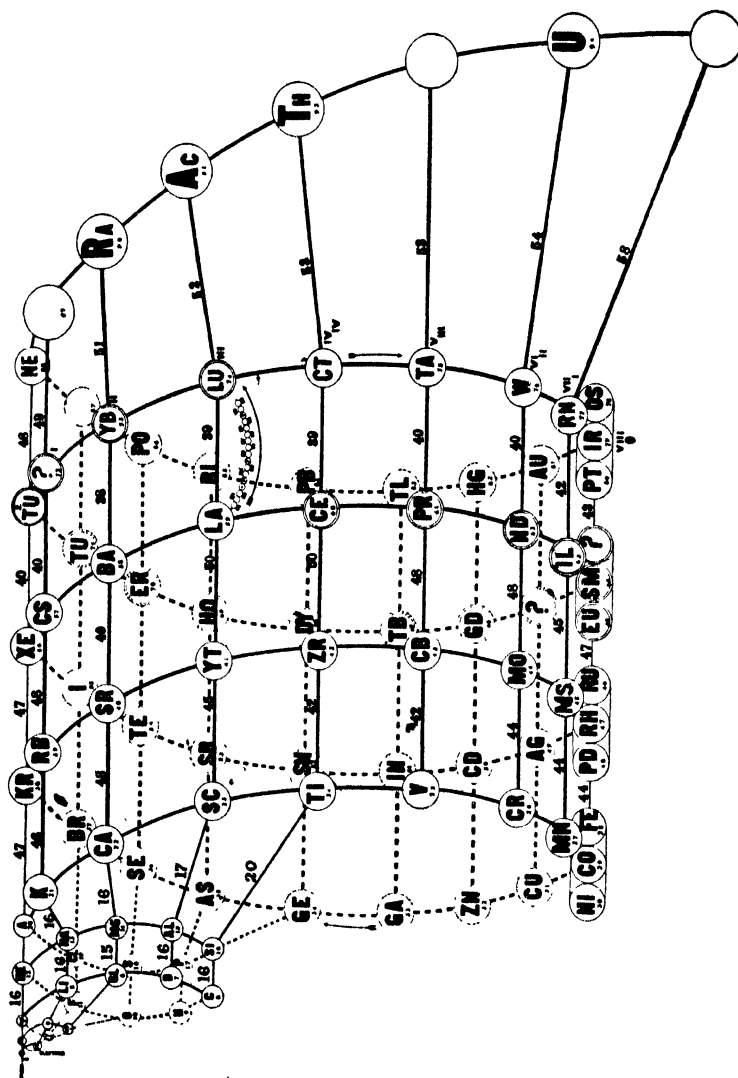


FIG. 1. A SIDE VIEW OF THE HELIX AS ORIGINALLY PUBLISHED

The Rydberg ordinal numbers (greater by two than the Moseley atomic numbers) and the transverse interweights are given. E at the beginning should be AE.

first octave, a single group of four elements, hydrogen to helium, in a single circle and a first group of two elements, ether and coronium, in a half-circle. In the original helix it was recognized that the numbers of the rings of the helix, which symbolized the succession of the elements as well as the number of elements in each type of ring, increased as the powers of two.

It will be noted that the elements are placed symmetrically in the order of their increasing atomic weights upon the successive coils of the helix. The distance between successive elements on the helix, the interspaces, are made proportional to the successive increments in atomic weight, the interweights. Thus in the half-octave with four elements ($H = 1$ to $He = 4$) the interweight is one and the interspaces are given this unit length. In the octaves the elements increase by two units of atomic weight, which are so represented as spaces. They average three each in the double octaves and four in the quadruple octave. As shown in figure 3, these interweights sometimes vary from the average by a full unit but in the figures the average is taken, which brings each group of common valence into a straight line. The use of the true values would make each line zigzag sharply yet remain separate from its neighbors.

THE EMENDED HELIX

In the original helix the ether and the electron were blended in a single symbol, E, which was joined with the element coronium in the first half-circle. This seemed infelicitous even then and Nicholson's investigations, demanding an additional element before hydrogen, caused a reconsideration of the question. It was then seen that the algebraic decrease in the powers of 2 and the geometric decrease in the coils of the helix had not been fully appreciated. The complete series of the elements can, on this plan of symmetry, be symbolized by the complete series of the powers of 2, i.e. $2^{-\infty} = 0; 2^{-1} = \frac{1}{2}; 2^0 = 1; 2^1 = 2; 2^2 = 4; 2^3 = 8; 2^4 = 16; 2^5 = 32$. The full geometrical implications of this series are shown in the helix of figures 2, 3, and 4. Thus if 2^2 represents a half-octave with four elements and 2^1 represents a fourth-octave with two elements, the question lies in the signifi-

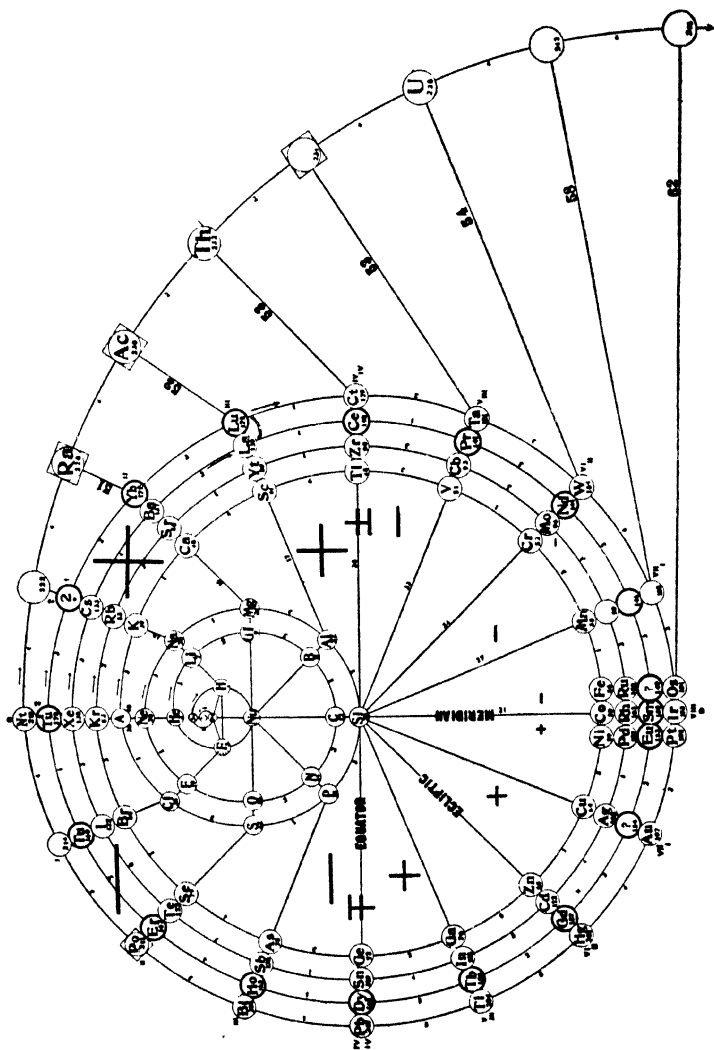


FIG. 2. THE HELIX FROM THE END

The longitudinal interweights are given and also the transverse interweights at the flares. The rare earths are each enclosed in a double circle and the short-circuiting from La to Lu is indicated by a long curved arrow. On the eighth line down, the symbol directly under He should be Cn instead of Si as printed.

cance of 2^0 , 2^{-1} , 2^{-2} , etc. Geometrically the case is one of decreasing radii, since a coil of the helix with four symmetrical radii makes place for four elements while the next smaller coil, with two radii, makes place for two.

The origin, O. The first step in the construction of the helix as a working model of the evolution of the elements is the fixation

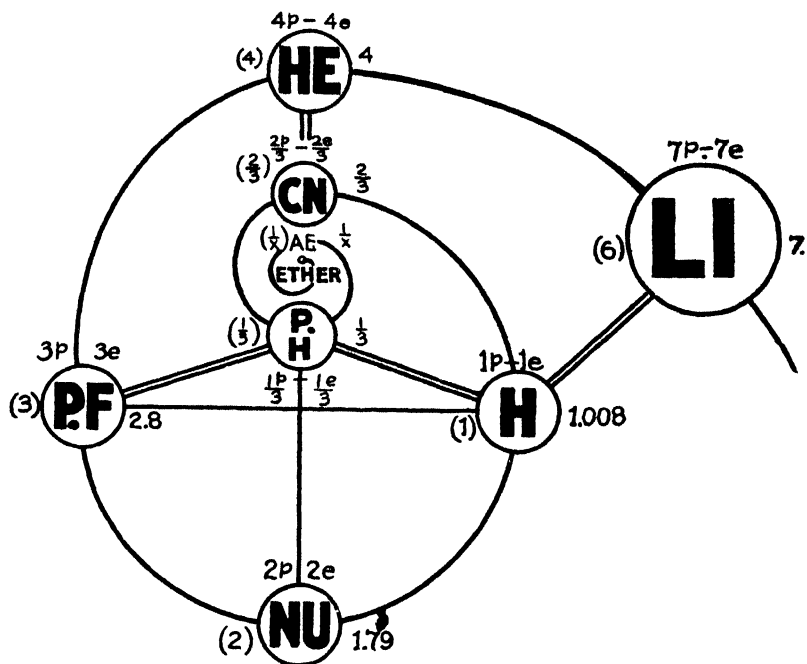


FIG. 3. ENLARGEMENT OF THE FIRST THREE COILS

Data given are the atomic composition, the ideal atomic weights in brackets and the observed or calculated atomic weights.

of the point of origin, O, in the ether or the electrical source of the vortex motion that is to form the simplest unknown elemental matter. This point, as being without dimensions, may be expressed by zero and, as it is a sample of every point in the infinite ether, its summation may equal infinity. The word ether is therefore placed at the beginning of the helix. Algebraically this is expressed by $2^{-\infty} = 0$.

The first circle; the ante-element. A radius vector occupies this point, O, and is at first without length, *et crescit eundo*. At the end of its first full rotation it has the infinitesimal length, O — AE (fig. 4), suggested by the mass of the primal substance which is formed by the vortex motion in the ether and typified by this rotation. This ante-element, AE, may thus represent the passage of the ether into the simplest transitional form of matter or the unknown action by which electricity passes into its first stable form. This may represent the electron, or, possibly, Ehrenhaft's sub-electron. The whole length of this circle represents the first interweight and is almost infinitesimal in all its parts.

The second circle; proto-hydrogen and coronium. Continuing the helix into the second circle, the two radii, O — PH and O — Cn, placed at the middle and the end of this circle, mark the positions of two undiscovered elements with fractional weights, named proto-hydrogen and coronium, which are thus at the ends of a single vertical diameter and represent algebraically the next distinctive term in the series $2^1 = 2$. These elements may be assigned one and two protons, respectively, as was done by Nicholson (2), and as is suggested by the equations for the elements derived from Thompson's formula for matter. All chemical work, however, is based on $H = 1$, so that $PH = \frac{1}{3}p - \frac{1}{3}e$, and $Cn = \frac{2}{3}p - \frac{2}{3}e$. These fractional values, while maintaining the proper ratio, may fitly typify these possible inert, transitional, coronal and nebular elements. "In fact," says Nicholson, "the nebulae must be composed of a set of chemical forms which are the simplest forms in which matter can exist. They are not isotopes because they have different spectra. The spectrum of a nebula may be described as the spectrum of chaos. Whatever may occur with terrestrial atoms, the electrons in a nebula are not held very firmly in the atoms and a continual interchange of electrons must be taking place, with a necessary bombardment of atoms by free electrons, to which the luminosity is probably due. The first products of an evolution of some form from these systems leads ultimately to the Wolf-Rayet stars in which series analogous to the terrestrial series may be found." The Wolf-Rayet stars are supposed to be developing into normal stars by the growth of

a simple nucleus at their centers. W. H. Wright (3) suggests that the nuclei of the planetary nebulae are closely related to the Wolf-Rayet stars and in many cases are such stars. He supposes that the order of condensation or settling into the nucleus from the nebulae is helium, nebulium and hydrogen.

On the helix proto-hydrogen is the first of the inert gases, standing above the center of the figure at the beginning of the series and symmetrically between hydrogen with a valence of $+1$ and proto-fluorine with a valence of -1 . Its valence is thus 0. Coronium, above proto-hydrogen, stands also in the series of the inert gases and so has a valence of 0. This second circle of the helix, with only one diameter, is not yet completely symmetrical and thus characterizes the two fractional elements. This is the first introduction of "matter."

The third circle. The next circle, the half-octave, is cut geometrically by four radii, typified algebraically by $2^2 = 4$. There is thus place for four elements and the previous interspace of $\frac{1}{2}$ becomes 1. These elements have respectively 1, 2, 3 and 4 electrons or protons and are formed by the successive addition of a single hydrogen atom. Their ordinal numbers (N) are the same as the number of protons contained and their atomic weights are approximately 1, 2, 3, and 4. This circle, with two diameters at right angles to each other, is the first completely symmetrical circle and so has the first stable and integral elements: hydrogen, the building stone of this first circle, and helium, the building stone of the last circle. These are the first terrestrial elements. They are accepted by Rydberg (see below) and, as possibilities, by Aston.

The fourth group of two circles. With the development of these two stable forms complexity becomes possible. In the next group of two circles ($2^1 = 2$) two hydrogen atoms tend to be added to each new form and so the interspace becomes two. Four diameters are drawn so that the number of interspaces becomes eight (algebraically $2^3 = 8$). Hence so stable a combination, the octave, has been reached that two circles are formed, the most typical, symmetrical and equally differentiated of the whole series. At the bottom of these two circles, at the very

center and umbilicus of the helix stand silicon and carbon, the ideal elements. One is the basis of the inorganic world and the other of the multitudinous compounds of the organic world. Both are embryonic triads.

The fifth group of four circles. Next, with the serial addition of 3H, a double-octave circle is formed having 8 diameters and thus place for 16 elements ($2^4 = 16$) which at first simulates the preceding circle but diverges downward from it into a larger circle. Passing titanium it goes down with increasing atomic weight, valence and density to the iron triad and rises with a symmetrical decrease of all these properties except atomic weight to germanium, which balances titanium opposite silicon. Thence it corresponds with the left half of the octave, reaching zero valence in krypton. It thus forms a circle of double size which occurs four times ($2^2 = 4$) with decreasing fidelity to type because of increasing bulk.

The triads. The equation, $2^4 = 16$, demands a circle of sixteen places. The three complete double-octaves which occupy these circles each contain 18 elements. The lowest position on each circle is therefore occupied by three similar elements, each having a valence of 8,—a triad. In completing the imperfect third double-octave, where three elements had to be supplied the most probable additions are indicated by question marks in the figures. One is the first element of a similar triad and another is the first element following the triad.

The sixth circle. Then comes the successive addition of 4H (or He), forming again a circle of double the previous size, a quadruple-octave. This becomes too complex for complete development or stability, as is indicated by the sudden appearance of radioactivity. Here the interspace becomes 4, the number of diameters 16, the number of elements 32, and the number of rings, if developed, would be 8. These numbers are extrapolations by the use of the series of the powers of 2 which have been found to apply in the earlier parts of the helix among existing elements.

The flares of the helix. If we subtract the atomic weights of the inner octave from the corresponding members of the outer

octave, as $C = 12$ from $Si = 28$, or $O = 16$ from $S = 32$, we get a constant difference of 16 because we are subtracting values which increase by two from values which increase by two. But if, in a similar manner, we subtract values in the outer octave from corresponding values in the first double-octave there is a gradual increase in the difference, from 16 to 32, since we are subtracting numbers which increase by two from numbers which increase by three. Following this flare the differences between the weights of corresponding members of the double-octaves again become constant at 48. But a new flare appears beyond the fourth double-octave and the differences again increase. These values may form a criterion for the judgment of those arrangements of the periodic table which, like that of Rydberg (4) adopted by Harkins (5), postulate two octaves, two double-octaves and two octo-octaves with 64 elements each. This demands a flare after passing xenon, but $Kr - Xe = 47$, $Sr - Ba = 49$, $Cb - Pr = 48$, $Mo - Nd = 48$, and so on, which shows that the third double-octave is strictly parallel to the two preceding ones. The arrangement of Hackh (6), which unites the third and fourth double-octaves in a single coil, is subject to the same criticism.

THE ATOMIC OR ORDINAL NUMBER; THE MOSELEY NUMBER

The number two when it stands for the nuclear charge of helium, its number of nuclear or of exterior electrons, or for half its atomic weight, represents fundamental characteristics of its atomic structure. When it is also used as its ordinal number or "atomic number" it expresses only the coincidence that helium happens to be the second *terrestrial* element recognized in the long established Mendeleeff table of the elements, which ignores the celestial elements after and perhaps before hydrogen. Throughout the two octaves there is an average increase of 2 in weight for each additional element so that the atomic weight is close to twice the atomic number. The heavier elements, however, have weights which are greater than twice their atomic number. The divergence begins at argon, i.e. where the double-octaves attach to the octaves. Thereafter the increment in weight is three and the

division of the atomic weight by two no longer gives the atomic number. The weights are, however, so large that the percentage error is small and the error is overlooked in the common statement of the relationship.

The Moseley series of atomic numbers does not exclude the possibility of further elements below helium. Nicholson (7) says, "It is not necessary to admit that the atomic number of an element corresponds *exactly* to its position in the table as we know it. . . . Other elements may exist between hydrogen and lithium besides helium, . . . and . . . the supposition that the atomic numbers of lithium, glucinum, and boron are (necessarily) 3, 4 and 5, respectively, cannot be admitted on theoretical considerations." Uhler (8), on the basis of the determinations of the x-ray spectra of the elements, declares that "the law must not be taken so literally as to mean that the x-ray data used by Moseley alone necessitate associating the integers 1, 2, . . . 13, . . . 74, . . . 92, with H, He, Al, W, and U, respectively." Indeed the atomic numbers need not even be integral.

Langmuir's concept of an octet of eight electrons has become established. A quartet is also possible and is suggested by symmetry for the lighter elements. Lande gives an elaborate mathematical justification for the arrangement of the centers of rotation of the electrons at the corners of a tetrahedron as well as at the corners of a cube. This model has been adopted, at least in part, by Langmuir (9).

THE RYDBERG ARRANGEMENT OF THE ELEMENTS COMPARED WITH THE HELIX

The Rydberg diagram (fig. 4) (11) offers a significant comparison with the helix. For this purpose the projection of the helix is represented in figure 5. The differences between the two are tabulated in table 1. The two differ chiefly in that Rydberg's table has two rings in each group while the number of rings in the successive groups of the helix increases in number as the powers of 2. Table 1 shows that the two differ immediately after the origin to permit the insertion of the two series representing 2^0 and

2¹ in the helix. Again, from the point of view of the helix the elements from hydrogen to helium form a single circle and not two loops, though this may remain somewhat uncertain until the valence of the intervening elements is known. From this point on the two tables are identical to the beginning of the fourth group. At that point the divergence of the two is radical, for the helix demands two additional double-octaves, or four in all, and utilizes the elements of the first loop of Rydberg's Group IV for this purpose.

In figure 4 the fourth group, beginning at xenon, should extend to the right to twice the width of the table but it is bent back at samarium in order to put homologous elements beneath each other, particularly the triads, Os, Ir, Pt, and Fe, Co, Ni. This forces the Rydberg diagram into a resemblance to the helix and compels the apparent insertion of an incomplete third double-octave of the rare earths, Cs to Tu, which is frankly recognized in the helix as an actual double-octave. This has been proposed by Bohr (12) in his discussion and revision of Rydberg's table.

In the helix provision for eighteen elements in each double-octave is accomplished by the triads in the octo-valent position. To provide the 18 elements in the rare earth series four unknown elements are inserted including three adjoining samarium and the element Tu 2, which is already widely accepted (13). Vogel (14) cites three separate culminations of chemical and physical properties at samarium and europium, which correspond with a triad culmination in the helix. He presents several considerations which favor a third double-octave for the rare earth group, as in the helix. Kirchof (15) has shown that the colored ions, which are found two or three places on either side of the triads in the normal double-octaves, are even more widely spread in the corresponding positions in the rare earths.

It would seem preferable to consider the thulium-samarium group as an abnormal double-octave among the others than to make it the beginning and type of a still larger group, as Rydberg does, because this latter would demand a marked similarity between the rare earths and the radio-elements which stand at the beginning of the second of these longest groups. But the dissimi-

	+7	+6	+5	+4	+3	+2	+1	0
G ₁	H(1) PF(3) Nu 2 He(4)							
G ₂	Li(5) — Be(6) — C(8) F(11) — O(10) — N(9)							
	Ne(12) Na(13) — Mg(14) — Al(15) Cl(19) — S(18) — P(17) Si(16)							
	Ar(20) K(21) — Ca(22) — Sc(23) — Ti(24) — V(25) — Cr(26) — Mn(27) — Fe(28)							Co(29)
	Kr(38) Br(37) — Se(36) — As(35) — Ge(34) — Ga(33) — Zn(32) — Cu(31) — Ni(30)							Rh(47)
	Rb(39) — Sr(40) — Y(41) — Zr(42) — Nb(43) — Mo(44) — — (45) — Ru(46)							Pd(48)
G ₃	X(56) I(55) — Te(54) — Sb(53) — Sn(52) — In(51) — Cd(50) — Ag(49) — ? Cs(57) — Ba(58) — La(59) — Ce(60) — Pr(61) — Nd(62) — — (63) — ? Tu(71) — Er(70) — Ho(69) — Dy(68) — Tb(67) — Gd(66) — — ? Lu(73) — Ct(74) — Ta(75) — W(76) — — (77) — Os(78)							Sm(64)
	Tu 2 Yb(72) — — (86) — Bi(85) — Pb(84) — Tl(83) — Hg(82) — Au(81) — Pt(80) — Ir(79)							
G ₄	Nt(88) Ra(90) — — (89) — — (87) — — (86) — — (85) — — (84) — — (83) — — (93) — — U(94)							

FIG. 5. MODIFICATION OF THE RYDBERG DIAGRAM TO CONFORM WITH THE HELIX

larity between them is striking and in the helix the most marked of all the flares separates the two groups and argues against any direct similarity.

The short-circuiting of the elements between lanthanum and tantalum as has often been proposed and is expressed in the

TABLE 1
Comparison of the Helix with the Rydberg Table

THE HELIX			THE RYDBERG TABLE		
Subdivision	Scheme	Number of elements	Subdivision	Scheme	Number of elements
Origin	$2^{-\infty}$	0	Group O 1st loop 2nd loop	2×0^2 2×0^2	$\left. \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right\} 0 = 0^2$
1 Circle	2^0	1 (ante-element)			
1 Circle	2^1	2 (fractional elements)			
1 Circle	2^2	4 ($\frac{1}{2}$ octave)	Group I 1st loop 2nd loop	2×1^2 2×1^2	$\left. \begin{smallmatrix} 2 \\ 2 \end{smallmatrix} \right\} 4 = 2^2$
2 Circles	2^3 2^3	$\left. \begin{smallmatrix} 8 \\ 8 \end{smallmatrix} \right\} (2 \text{ octaves})$	Group II 1st loop 2nd loop	2×2^2 2×2^2	$\left. \begin{smallmatrix} 8 \\ 8 \end{smallmatrix} \right\} 16 = 4^2$
4 Circles	2^4 2^4 2^4 2^4	$\left. \begin{smallmatrix} 16 \\ 16 \\ 16 \\ 16 \end{smallmatrix} \right\} (+2 \text{ in triads})$ (4 double-octaves, total, 72)	Group III 1st loop 2nd loop	2×3^2 2×3^2	$\left. \begin{smallmatrix} 18 \\ 18 \end{smallmatrix} \right\} 36 = 6^2$
8 Circles	2^5	32 (quadruple-octaves)	Group IV 1st loop 2nd loop	2×4^2 2×4^2	$\left. \begin{smallmatrix} 32 \\ 32 \end{smallmatrix} \right\} 64 = 8^2$

(Only 6 radioactive elements formed in first circle or loop)

diagrams of Soddy and of Harkins, is a practical convenience but it is a surrender rather than a solution. It is expressed in the figures 1 and 2 by an arrow joining La with the main curve below Ct and by indicating the omitted ring by heavier circles around the symbols of the elements omitted. Rydberg rejects entirely

this short-circuiting plan and has made elaborate calculations of the positions of the rare earth elements. His arrangement of the individual elements has been closely followed in the helix.

REFERENCES

- (1) Amer. Chem. J., **45**, 160 (1911).
Science, **34**, 160 (1911).
- (2) Monthly Not. Royal Astron. Soc., **74**, 506 (1914).
Harkins and Wilson, J. Amer. Chem. Soc., **37**, 1410 (1915).
- (3) Astrophysical J., **40**, 466 (1914)
- (4) Lunds Universitets Arsskrift, Bd IX, No. 18 (1916).
- (5) J. Amer. Chem. Soc., **38**, 218 (1916).
- (6) Das Synthetische System der Atome. Hephaestos Verlag. Hamburg, 1914, p. 26.
- (7) Phil. Mag., **27**, 540 (1914).
- (8) Proc. Nat. Acad. Sci., **3**, 89 (1917).
- (9) Science, **51**, 605; **52**, 433 (1920).
- (10) The Magnetron Theory of the Atom. Smithsonian Miscellaneous Collections, **63**, No. 11 (1915).
- (11) Phil. Mag., **28**, 139 (1914).
- (12) N. Bohr, Theory of Spectra and Atomic Constitution. N. Y., Macmillan Co. (1922.) Fig. 1, p. 70
Kramers and Holtz, The Atom and the Bohr Theory. N. Y., A. A. Knopf. (1923.) Fig. 34, p. 201.
- (13) S. I. Levy, The Rare Earths N. Y., Longman, Green & Co. (1924) p. 231.
- (14) Z. Anorg. Chem., **102**, 175 (1918).
- (15) Physikalische Z., **21**, 717 (1920).

ACID AND BASIC CATALYSIS

J. N. BRÖNSTED

University of Copenhagen, Denmark

TRANSLATED BY KARL H. SANDVED AND VICTOR K. LAMER

Columbia University, New York, New York

CONTENTS

Preface	231
Introduction	231
I. The influence of solvation in catalysis by hydrons	234
II. The dual theory of catalysis.	245
III. The ionization theory of catalysis.	252
IV. Primary salt effect, activity and concentration.	265
V. Secondary salt effect	278
VI. The acid-basic function.	284
VII. On the strength of acids and bases.	291
VIII. An extended theory of acid and basic catalysis.	312
IX. Appendix.	334
References	330

PREFACE

The present review is a translation of a monograph: "*Om Syre- og Base-katalyse*" published by the University of Copenhagen September, 1926. The field is covered in the text up to March, 1926.

In order to bring the article up to date a short *Appendix* has been added in which the progress in 1926-1927 has been presented.

The author is much indebted to Mr. K. H. Sandved and Professor V. K. La Mer for the trouble they have taken in making the translation.

INTRODUCTION

Acid and basic catalysis constitutes an important chapter in the history of physical chemistry. From studies in this field originated the chief contributions to our knowledge of homogeneous catalysis. In the formulation of the laws governing the

velocity of chemical reactions in general, investigations of acid and basic catalysis have played a conspicuous part as appears from the classical investigations of Wilhelmy (180) on the inversion of cane sugar by acids and from the various catalytic studies by Ostwald (149) and Urech (176). Furthermore consideration of the phenomena of acid and basic catalysis has contributed to the understanding of the mechanism of chemical reactions in general and also has in many cases been useful in the verification of Arrhenius' theory of electrolytic dissociation (8).

The ideas of acid and basic catalysis, developed by Arrhenius (9, 10, 11, 12) on the basis of his dissociation theory, are still commonly recognized as principally correct. The chemist still considers the phenomena of acid and basic catalysis as due to the effect of hydrogen and hydroxyl ions respectively and makes use in numerous ways of such a dependence as a means of determining the concentration of these ions.

This simple conception of catalytic phenomena to which the investigations of Arrhenius have led, however, failed to prove entirely satisfactory. First, as is natural, certain factors used in the theory have on closer examination revealed incompleteness of definition. Secondly it must be admitted that the difficulties, encountered from the very beginning in the quantitative working out of the theory of electrolytic dissociation, have presented themselves with especial emphasis in the field of reaction velocity. Arrhenius (11) himself called attention to the difficulties caused by such anomalies in the consideration of catalytic phenomena from the point of view of electrolytic dissociation: thus for the interpretation of the effect of acids on the inversion of cane sugar he found it necessary to assume an increase in the "activity" of the sugar molecules on addition of the acid besides the ordinary catalytic effect of the hydrions. Later on as a result of similar difficulties he resorted to formulating the velocity law by means of osmotic pressure instead of concentration.

Other authors, engaged in the field of catalytic phenomena, have felt the same difficulties in different ways. In the course of time numerous attempts have been made to effect a closer approach to the experimental data by introducing suitable modifica-

tions in the original theory. While most of these attempts have failed in presenting decisive results, they have brought about a collection of data, which should be of value in future work. The influence on the catalytic velocity of reaction produced by addition of salts or more generally by a change in the concentration of electrolytes forms the chief basis for these modifying theories. It is primarily this conception called "Neutral salt effect" or more correctly "Salt effect," that has given rise to difficulties in the interpretation of the kinetic phenomena of reaction. In some respects these difficulties when closely studied are quite imaginary. In other respects, however, points of great kinetic significance are attached to the term "Salt effect," the understanding of which has been rendered possible only by means of the new theory of ionic solutions, which stands out as one of the chief advances in the recent development of modern chemistry.

In addition, views of interest for acid and basic catalysis have been developed in other directions. This applies to the ideas advanced on the state of the catalyst in solution and to the theory ascribing catalysis to the ionization of the substrate or to the formation of an intermediate product which reacts spontaneously.

One of the aims of this review is to give a critical exposition of the theories of acid and basic catalysis, without in any way laying claim to having attained a complete survey of this intricate subject. The intention is furthermore to give a brief outline of the modern views on catalytic salt effect. Finally there will be given certain ideas and considerations concerning the phenomena of acid and basic catalysis based upon views of the acid-basic function which have recently been advanced and for the development of which the catalytic decomposition of nitramide and certain analogous reactions like the mutarotation of glucose have furnished the experimental foundation.

The much disputed question as to a general definition of the concepts of catalysis and catalyst will not be treated in this paper, the laws governing the catalytic processes discussed here being considered as quite identical to those valid for other homogeneous reactions. Thus for example the catalytic decomposition of cane sugar by hydriions will be considered as a bimolecular

reaction between the sugar molecule and the hydrion, for which the differential equation for bimolecular reactions is supposed to be applicable to the same extent as for non-catalytic reactions. The fact that one of the reacting molecules, viz. the catalyst, is to be found among the products of the reaction according to our view cannot possibly affect the laws governing the process.

I. THE INFLUENCE OF SOLVATION IN CATALYSIS BY HYDRIONS

Arrhenius' theory of acid catalysis as a catalysis caused by hydrions does not distinguish between the hydrion in the free and in the hydrated state. When we realize that the water-free and the hydrated hydrion in reality are two entirely different molecules, the question then arises as to their relative rôle in the hydrion catalysis. Resulting differences in the effects of the two ion species might perhaps offer a possibility of interpreting some of the anomalies, brought about by addition of salt or by some other change in the solvent.

Lapworth (66, 122) was the first to subject this question to a thorough treatment. According to his theory, which in many ways may seem plausible, the catalytic effect in acid catalysis is due solely to the non-hydrated or more generally to the non-solvated hydrion. The solvated hydrion is supposed to be inactive. Nowadays it is well known, that the free hydrion—an electron-free nucleus—is altogether different from any other chemical molecule, and it might therefore appear well founded to attribute special catalytic properties to the hydrion. On the other hand this would give the hydrion an unique position which is not supported by experimental evidence. The hydroxyl ion for instance, is equally as effective in catalytic processes as the hydrion.

The most important experimental foundation on which Lapworth's theory is developed comprises the results obtained by studying the effect of small quantities of water on saponification and other reactions catalyzed by hydrions, for example the hydrazobenzene-benzidine arrangement and the bromination of ketones in non-aqueous solutions. About 1895 H. Goldschmidt (72, 73) showed that hydrion catalysis in alcoholic solution is

considerably depressed on addition of water. In subsequent papers he has studied the phenomenon thoroughly and developed a theory, which will be extensively dealt with below. Other authors have also (179, 105, 24, 175, 50) examined the influence exerted by water on various reactions catalyzed by hydrions in alcoholic and related solutions, and in this way have in the main confirmed the existence of the anti-catalytic effect of very small concentrations of water.

Lapworth explains the effect by the assumption that most of the hydrions in aqueous solution are hydrated having taken up a molecule of water to form the ion H_3O^+ . In accordance with recent authors this ion will be termed here the oxonium ion, its great importance in the theory of acid and basic catalysis being further developed in chapters 6, 7 and 8. The free hydrion is, however, supposed to exist in the solution in small quantities, according to the following equilibrium:



In an alcoholic solution for instance, an increase in the concentration of water will diminish the quantity of free hydrions, these being now transformed into oxonium ions. The depression of the catalytic action will, according to Lapworth, be naturally explained as due to a decrease in concentration of the effective catalyst.

The bromination of acetone may be taken as an example. The partial reaction which is measurable kinetically and which is strongly catalyzed by hydrions, consists most probably in the enolization (121) of the acetone, the depression caused by addition of water being about proportional to the quantity added. Only 2 per cent of water is necessary for reducing the velocity to 1/10 of the original value in a water free solvent.

Lapworth tries to support his views on the free hydrion by some other methods as well. For this purpose Hardman and Lapworth (92) have carried out a series of E.M.F. measurements in cells of the following type:



in which the quantity of water is small as compared to the quantity of alcohol, but great as compared to that of hydrochloric acid. The E.M.F. of such a cell is given by the formula:

$$E = RT \ln \frac{P}{P_0}$$

where P is what the author calls the "availability of the acid," a term which evidently is identical with what we now would call the activity of the free hydron. The measurements conform with the expression:

$$\frac{P}{P_0} = \frac{r}{r + w} \quad (1)$$

where r is a constant and w denotes the molarity of water in the alcoholic solution.

In reality a formula in agreement with (1) can easily be derived from the classical law of mass action. By using the following three equations

$$\frac{c_{H^+} c_{C_2H_5OH}}{c_{C_2H_5OH_2^+}} = K_1,$$

$$\frac{c_{H^+} c_{H_2O}}{c_{H_3O^+}} = K_2,$$

and

$$c_{C_2H_5OH_2^+} + c_{H_3O^+} = c_{H^+}$$

where c_{H^+} is the total concentration of acid, assuming complete dissociation we get:

$$\frac{c_{H^+} \text{ (alcohol + water)}}{c_{H^+} \text{ (alcohol)}} = \frac{\frac{c_{C_2H_5OH}}{K_1}}{\frac{c_{C_2H_5OH}}{K_1} + \frac{c_{H_2O}}{K_2}} = \frac{r}{r + c_{H_2O}}.$$

According to the terminology proposed in Chapter 7, r is the conventional dissociation constant in alcoholic solutions. The hypothesis of the existence of the free hydron is consequently not

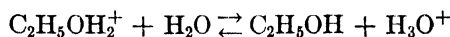
in conflict with the results of the E.M.F. measurements. Lapworth considers this to be a support of the theory attaching importance especially to the fact (124) that the two catalytic constants k and k_0 for different catalytic reactions with and without addition of water respectively follow a formula analogous to (1):

$$\frac{k}{k_0} = \frac{r}{r + w} \quad (2)$$

r has here approximately the same value as given by the E.M.F. measurements.

Lapworth's theory has been given much credit and has been supported chiefly by English authors (42, 46). It will be shown, however, that there are decisive difficulties opposed to the theory.

In their first paper Fitzgerald and Lapworth call attention to the conflict between their views and those of Goldschmidt, without making any effort to test the relative soundness of the two theories. Goldschmidt, who as before mentioned, was the first to observe the anti-catalytic effect of water, explains the phenomenon by ascribing the catalytic effect to the hydrogen ion-alcoholate (83, 74, 80). The effect of the water depends consequently upon a partial decomposition of the alcoholate converting the hydrion into the oxonium ion according to the scheme:



In this system the alcohol serves as the solvent and hence may be considered to be present in constant concentration. The mass action law applied to this system yields:

$$\frac{c_{\text{C}_2\text{H}_5\text{OH}_2^+} c_{\text{H}_2\text{O}}}{c_{\text{H}_3\text{O}^+}} = r,$$

where r as previously mentioned corresponds to the dissociation constant of the oxonium ion in alcoholic solutions. If C_{H^+} means the total hydrion concentration regardless of solvation, the following equation is easily derived:

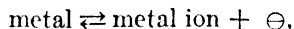
$$\frac{c_{\text{C}_2\text{H}_5\text{OH}_2^+}}{C_{\text{H}^+}} = \frac{r}{r + w}$$

If the hydrion alcoholate functions as the catalyst this equation may be transformed into the following:

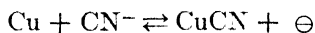
$$\frac{k}{k_0} = \frac{r}{r + c_{H_2O}}$$

k and k_0 being the velocity constants in alcoholic solutions with and without water respectively. The assumption of the hydrion alcoholate as the active catalyst then leads to the same formula for the anti-catalytic effect as does the assumption of the catalytic activity of the free hydrion. This is also a simple consequence of the fact that the hydrion alcoholate and the free hydrion must be present in the alcoholic solution in a constant ratio with respect to their concentrations. The existing data cannot therefore be interpreted as verifying one of these theories at the expense of the other; they both explain equally well the experimental facts.

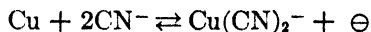
Furthermore it should be noted that the E.M.F. measurements mentioned above can not possibly prove anything as to the real existence of the free hydrion. The idea is frequently met that the validity of Nernst's formula (145) for the electromotive force, from which formula (1) is derived virtually presupposes a real concentration of the ion under consideration. However, a well defined chemical potential, for instance of a metal in a salt solution, may be obtained, even when no appreciable concentration of the metal ions exists in the solution. In the numerous instances in which Nernst's formula leads to absurd ion concentration—as for instance in the case of a copper electrode in a potassium cyanide solution—it is therefore unjustified to consider the reaction



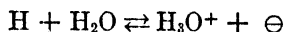
where \ominus denotes the electron, as always determining kinetically the potential at the electrode. Under such circumstances and also in general the potential between electrode and electrolyte might be defined by means of more indirect reactions. For the copper-copper-cyanide system we might for instance have:



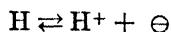
or



As regards the hydrion potential in presence of water the indirect kinetic reaction:

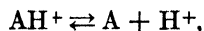


can just as well as the direct reaction:



account for the existence of a well defined potential at the electrode.

The definition of the potential cannot therefore give any information as to the real existence of the water-free hydrion. Nor can the agreement between the electromotive force and the catalytic measurements from Lapworth's point of view be of decisive importance, as may also be seen from the thermodynamic scheme of equilibrium:

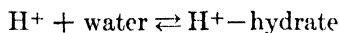


where *A* is the alcohol molecule, or more generally, the molecule of the solvent employed. According to this scheme the concentration of the hydrion alcoholate is proportional to the concentration calculated for the free hydrion. The relation between catalytic activity and electromotive force, which is derived if the free hydrion is supposed to be the active molecule in the catalysis as well as in the creation of the electromotive force, will consequently be maintained if the hydrion alcoholate be considered active in both respects.

It is thus impossible on this basis to choose between Lapworth's and Goldschmidt's theories, although other considerations might be taken into account to show the doubtful character of ascribing any kinetic significance to the free hydrogen nucleus. It is hard to believe that the hydrogen nucleus with one positive charge and without protecting electrons should be capable of existence amidst other molecules possessing electron systems acting as screens against the surroundings. This impossibility is according to

Fajans (62) manifested by the enormous heat effect of 232,000 cal. produced when the free hydrion unites with water. If this heat quantity is supposed to represent the affinity of the hydrion for water—as is approximately correct, the concentration of the free hydrion in an acid solution will be calculated equal to 10^{-180} . This calculation is based upon an application of the gas laws which naturally break down at these concentrations, but no doubt the order of magnitude of the figure demonstrates that the free hydrion cannot be present in aqueous solutions at kinetically significant concentration. The catalytic effect must consequently be assigned to the hydrion hydrate or the oxonium-ion, while in alcoholic solutions the hydrion alcoholate should be considered as the active catalyst in agreement with the view put forward by Goldschmidt.

Lapworth's theory on acid catalysis has recently been taken up by F. O. Rice (154) in his work on "chemical reactivity." Rice's views are based chiefly upon considerations regarding the temperature coefficient of catalytic reactions for which he takes advantage of Arrhenius' (11) well known theory of the dependence of the velocity of reaction upon the temperature, in which the equilibrium between the molecules of the substrate in normal and "active" state is supposed to be displaced in favor of the active molecules with increased temperature. This view is applied by Rice to the catalyst, whereby the shift of the equilibrium:



is made to account for the dependence of the reaction upon variations in temperature. The solvent is not supposed to take part in the reaction and the whole catalytic effect is supposed to be exclusively due to the free hydrion.

The author presupposes (apparently as an obvious fact) that the temperature coefficient of all reactions belonging to this class must be the same provided the theory is true and he postulates, moreover, that the reactions catalyzed by hydrions and hydroxyl ions can be resolved into comparatively few classes. It is mentioned as a verification of the theory that the experimentally determined temperature coefficient $k_{35^\circ}/k_{25^\circ}$ in acid catalysis of

esters is found to be between 2.2 and 2.5, while this identity can be explained from the old theories only by ascribing the same heat of activation to all the reacting molecular species. The temperature coefficient for the simple non-hydrolytic reactions catalyzed by hydrions is according to Rice:

$$\frac{k_{35^\circ}}{k_{25^\circ}} = 3.08.$$

This theory of the free hydrion as the only important factor in acid catalysis is extended further in deriving more general conclusions: The chemical reactions, pictured by the usual stoichiometric schemes, are as a rule far from giving the real mechanism. Products of dissociation and chemical compounds with the solvent, although present in exceedingly small concentrations, may be the true participants in the reaction. Such an assumption suffices for explaining apparent reaction anomalies, for the interpretation of which it will be unnecessary, therefore, to ascribe a kinetic effect to the surrounding medium.

In addition to what has been said before of the probable concentration of the free hydrion in water solution, there are further objections to which this "theory of reactivity" can be exposed. The contention that the theory of the free hydrion requires the same temperature coefficient for all simple reactions catalyzed by hydrions, can hardly be maintained because in that case the temperature coefficient for all bimolecular reactions involving the hydrion and generally also for all reactions with one common molecule ought to be the same. The identity of the temperature coefficients, experimentally found, cannot therefore be considered as evidence in favor of Lapworth's theory. Nor can the assumption that all reactions catalyzed by hydrions possess temperature coefficients belonging to comparatively few groups be taken as supporting the theory.

The hydrolytic decompositions catalyzed by hydrions may be considered as a separate group, different from the non-hydrolytic reactions mentioned above. In order to explain why the coefficient falls from 3.08 to 2.4 on proceeding to this type of reaction, Rice assumes the process to take place between the free hydrion

and the hydrate of the substrate, at the same time pointing out that only a small quantity of the latter (dependent upon the temperature) is present in the hydrated state. This explanation will not fit, however, unless the decomposition of the hydrate varies in the same manner with the temperature for all hydrolytic decompositions, and this is again tantamount to the introduction of a constant heat of activation. While it is admitted that Rice, by making these assumptions, can arrive at the same temperature coefficients for all hydrolytic decompositions, the existence of the "comparatively" few groups of temperature coefficients has not by any means been proved in this way from theoretical views on acid catalysis.

The more general considerations made by Rice on the great difference between stoichiometric and kinetic conditions are no doubt quite sound. This question, which has been thoroughly treated by Kendall (109, 112, 110), has been generally recognized, although it might not always have been sufficiently respected. It ought to be emphasized, however, that the kinetic equations often are quite independent of whether the real equation or a thermodynamically equivalent equation is employed. (See section 3.) Be this as it may, it still remains unjustified to make the reactivity of the molecules independent of the medium in which the reaction proceeds.

Mention ought also to be made of the extension of the Lapworth-Rice theory of acid catalysis to account for the corresponding phenomena in basic catalysis. This extension made by Rice gives a comparison between the two kinds of catalysis. The hydroxyl ion is supposed to be almost completely hydrated, and the catalytic effect considered as due to the small amount of non-hydrated hydroxyl ions. Since the hydration tendency is much greater for the hydrion than for the hydroxyl ion the non-hydrated part will assume a higher value for the latter. The concentration of free hydrions in "neutral" solution will consequently be much lower than the concentration of the free hydroxyl ions although the hydrated ions are present in practically equal amounts.

There may be some reality in these qualitative considerations. Rice, however, makes the additional assumption that the real

catalysts in acid and basic catalysis are of the same efficiency. If the real catalysts were the hydrated hydrions and hydroxyl ions, acid and basic catalysis ought then to be the same at the "neutral point." A solution, being "stoichiometrically neutral," would also be "catalytically neutral." The fact that the reaction minimum, or the stability maximum (Euler), for many reactions lies beyond the neutral point, for example, close to $p_H = 5$, is explained on the basis of the above assumption of the free ions being the active catalysts, whose concentrations at the neutral point are supposed to be quite different. The two water free catalysts are present in equal concentrations at the point of reaction minimum or at the "catalytic neutral point" as it is called by Rice; the location of this point on the acid side being due to the greater hydration of the hydrion as compared to that of the hydroxyl ion.

These considerations are, however, largely arbitrary. The assumption that acid and basic catalysts are equally effective is not well founded and may rather be considered as highly improbable. The interpretation given of the reaction minimum must therefore be rejected as being theoretically unsound. It is moreover well known that no fixed location can be attributed to this point, although it seems to lie about $p_H = 5$ in many cases. A great number of reactions are recorded where the effect of hydrions and hydroxyl ions is altogether different. Nitramide is completely unstable in alkaline solutions, while quite insensitive to acids. The salts of nitrosohydroxyl aminesulphonic acid are decomposed instantaneously in acid solutions while being reasonably stable to alkalis. Numerous examples of this kind exist, and are quite incompatible with an assumption of the same catalytic effect for the hydrion as for the hydroxyl ion, whether water-free or in the hydrated state.

An objection of more formal character may be added. As mentioned above, Rice formulates the concept of "catalytic neutrality" in relation to that of "stoichiometric neutrality." Stoichiometric neutrality is neutrality in the sense used by the chemist to denote the state in which the concentrations of hydrions and hydroxyl ions attain the same value. As pointed

out elsewhere (30) it is intrinsically illegitimate, however, to make this conception of neutrality imply any sort of counterbalancing effects, such as in the case of electric neutrality. The reason is that ideas like acidity and basicity, strength of acid and strength of base do not have any common measure. Acidity can only be determined in relation to another acidity, basicity to another basicity. It is obvious that the acidity will decrease and the basicity increase, when sodium hydroxide is added to hydrochloric acid. It may be said that the acidity is diminished at the same rate as the basicity is increased, but the point where they become identical cannot be fixed, because this question has no logical sense.

There seems therefore no reason why the concept of stoichiometric neutrality should be generalized further into the field of kinetics. The common neutrality concept is based upon the idea of expressing equality of two effects of opposite kind, while the effects attaining the same value in the catalytic neutral point are entirely of the same kind. It is therefore advisable to omit the concept of catalytic neutrality.

Although this is a purely formal question, it ought not to be dismissed as immaterial. The possibilities for development in a theoretical field are often dependent to a considerable extent upon suitable definitions and logical conceptions, whereas natural progress can be hampered or led astray if the definitions are lacking in clarity and cogency.

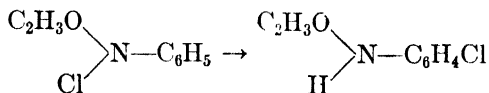
Nothing has been revealed in the theory of catalysis by adopting the theory of Lapworth, which can modify the result arrived at above as to the importance of solvation. Goldschmidt's views, in which the catalysis is ascribed to the hydrionhydrate in aqueous solutions, to the alcoholate in alcoholic solutions, to the anilinium ion (81) in aniline solutions etc. must still be considered to offer the best foundation for a theory of catalysis. This theory, so to speak, is imposed upon us by its own force, when we realize that the free hydrion is incapable of existence in any solvent and consequently either has to remain in the original acid molecule or to combine with the molecules of the solvent.

This mode of thinking leads to the assumption of different

"hydrions" in different solvents, so that hydron catalysis in different systems is really produced by different catalysts. This result is, however, as shown in Chapter 8, only the first step in the establishment of a more general theory of acid and basic catalysis.

II. THE DUAL THEORY OF CATALYSIS

The observations made by Blanksma (20) and later by Acree and Johnson (5), studying the transformation of acetyl chloro-amino benzene into p-chloracetanilide:



are among the first experimental data to have contributed to the formulation of the theory of catalysis later called the dual theory. This reaction is catalyzed by hydrochloric acid, but the velocity of reaction is proportional to the square of the acid concentration and not to the first power as in simple hydron catalysis. If the phenomenon be interpreted as an effect due to the HCl molecule, a second power law will be obtained, provided that no other substances add to the catalytic effect. The more general conclusion has been drawn that the hydron is not the only catalyst in acid catalysis, a greater or smaller effect being ascribed also to the undissociated acid molecules.

By studying the reaction mentioned above, many investigators (6, 148, 156, 96) have been able to show that the mechanism of the total reaction is more complicated and that the partial reaction determining the velocity is catalyzed by hydrions as well as by chloride-ions, which kinetically amounts to an HCl-catalysis. This reaction is not suitable, therefore, as a type of the ordinary hydron catalysis, being more related to such reactions as the decomposition of diazo acetic ester by HCl (25, 67) and the addition of HCl by quinone (50). These processes will not proceed unless chloride ions as well as hydrions are present.

Investigations on catalytic salt effect in alcoholic as well as in aqueous solutions have given the dual theory a more solid founda-

tion. On the whole, the development in theories on acid and basic catalysis might to a great extent be considered as a reflection of the interpretations given in the course of time for the influence of salt upon the velocity of reaction. The dual theory, as formulated by Senter (163) and Lapworth (124), has gained support from many sides, in fact it is represented in several text books (128, 181) as an unquestionable reality. It must be realized, however, that the results of the theory are due to wrong interpretations of experimental data.

Since the conditions in aqueous and alcoholic solutions are in a quantitative sense very different, it appears natural to deal first with the simpler aqueous solutions. The dual theory has been brought forth chiefly by H. S. Taylor (168, 169, 170, 171) and Dawson (44, 45) and his coworkers. Taylor worked especially on the hydrolysis of esters, catalyzed by strong or moderately strong acids with and without addition of neutral salts. By using hydrochloric acid as a catalyst and potassium chloride as neutral salt he thus found for five different esters the ratio between the catalytic constants in 0.1N HCl + 1.0N KCl and in 0.1N HCl without any KCl to be 1.24, the deviations from this value amounting only to a few per cent. 1N KCl consequently causes the velocity to rise 24 per cent. The ratio between the catalytic constant of the HCl-molecule and the hydrion can be calculated if the increase in undissociated HCl, brought about by addition of KCl and determined by conductivity measurements, is made to account for the effect observed. The experiments gave an average value 2.9 for this ratio. Other concentrations of the acid gave somewhat lower values, yielding the final average of 2.77, whereby the catalytic effect of the HCl-molecule would be about three times that of the hydrion.

It will be seen that this explanation necessitates the assumption that the salt effect would be zero if no increase in the concentration of the undissociated HCl had taken place on addition of KCl. This assumption, however, is absolutely without foundation. Changing the solution from 0 to 1N KCl means a very serious alteration of the medium kinetically. For instance Geffcken's (71) solubility measurements show that the activity coefficient of a

non-electrolyte like oxygen, is subject to a change of about 38 per cent on addition of NaCl up to 1N concentration. It seems plausible, therefore, to interpret the catalytic effect mentioned above as a pure kinetic effect, independent of the formation of new molecules. This explanation conforms much better to the modern view on electrolytes, in which the hydrochloric acid under these conditions is supposed to be completely dissociated.

A more direct corroboration of this explanation may moreover be found in another experimental series by Taylor. An almost constant ratio, varying from 1.24 to 1.20, will be found if the effect of 1N KCl on the catalysis by hydrochloric acid in concentrations varying from 0.01N to 0.5N is calculated on the assumption of constant hydrion concentrations. These values of the ratio will amount to 1.42 and 1.55 respectively, i.e., much more diverging figures if the hydrion concentration is supposed to be reduced by addition of KCl according to the classical method of calculation.

This is only a single example of the simplification brought about in the field of kinetics by employing the stoichiometric concentration of strong electrolytes instead of introducing conductivity corrections or activities. This last point will be thoroughly treated in Chapter 6.

The calculations by Snethlage (165) on Palmaer's (150) inversion experiments are subject to similar objections as were brought against those by Taylor. The inversion velocity caused by 0.1N KCl is about 6 per cent higher than what is calculated from the velocity at concentrations lower than 0.01N. The catalytic effect may now be calculated if this 6 per cent increase is supposed to be due to the increase in undissociated HCl. The salt effect may, however, just as well be interpreted kinetically. The latter view seems to be supported by the minor variations in the inversion constant at concentrations lying between 0.01N and 0.001N HCl, because the formation of catalytically active HCl within this concentration range would give a considerably higher variation than what is observed.

Taylor's calculations lead also in the case of tri- and dichloroacetic acid to a catalytic effect of the undissociated molecules, although much less than for HCl. The experimental error is much

greater for these weak electrolytes, however, while at the same time the application of the conductivity method and the fact that the primary and secondary salt effect are not taken into account make the results very problematic.

Results of somewhat greater significance for the dual theory have been obtained by H. M. Dawson (43, 44, 45) and his co-workers from investigations on enolization of acetone catalyzed by acids. The catalytic effect of certain organic acids was studied at various concentrations of acid, with and without addition of the corresponding salt. The results with monochloroacetic acid are of special interest. The ratio between the catalytic effect and the hydrion-concentration undergoes a considerable increase, when the concentration of the system with respect to either acid or salt is increased, this effect being much greater here than in the case of ester hydrolysis. The effect might be ascribed to the undissociated molecules, as assumed by Dawson. In order to enable any definite conclusions it is however absolutely necessary to take account also of the salt effect mentioned before, as well as of the influence of the monochloroacetic acid in changing the medium.

The catalytic values for stronger acids like trichloroacetic, trichlorobutyric and naphthalene sulphuric acids etc. are so uncertain that no importance can be attached to them, the reasons being the same as in the case of hydrochloric acid.

Dawson and Reimann (45) point out that their results do not agree with the investigations by Arrhenius (12) on the inversion of cane sugar with and without addition of the corresponding neutral salt. This is true also when the calculations are carried out according to more recent views. It seems as if the effect, for instance, of acetic acid on the inversion velocity is determined only by the hydrion concentration, if the various salt effects are taken into account. It must be considered, however, that acetic acid is much weaker than monochloroacetic acid, investigated by Dawson and Reimann, while also the nature of the catalyzed reactions certainly influences the law governing the process of catalysis. The importance of the latter point will be taken up in one of the following chapters.

The catalytic effect of an acid from the viewpoint of the dual theory may be expressed by the equation:

$$h = k_{H^+}c_{H^+} + k_M c_M,$$

where h is the catalytic velocity, k_{H^+} , k_M , c_{H^+} and c_M being the catalytic constants and concentrations of the hydrion and the undissociated acid respectively. Taylor (171) and Snethlage (164) have discussed the importance of the dissociation constant K of the acid, and the former has given the following generalization:

$$\frac{k_M}{k_{H^+}} = \sqrt{K}.$$

According to this equation k_M must be greater than k_{H^+} for strong acids. In certain cases the effect of the hydrion ought to equal that of the undissociated acid, and the catalysis would consequently be independent of the degree of dissociation (78). The evidence given in support of this equation is, however, very scant.

Numerous investigations on the dual theory (1) have been carried out in alcoholic solutions, especially by Goldschmidt (75, 82, 76, 77), Bredig (23), Snethlage (164, 165) and Acree (3). Many of the results obtained give apparently good reasons for the maintenance of the dual theory. The calculations are, however, based upon the classical mass action law, and they generally assume the complete absence of kinetic salt effects. In both of these respects the alcoholic solutions deviate more from ideality than do water solutions, and therefore the results obtained on this basis must be considered very critically.

In the development of his theory on strong electrolytes Bjerrum (17) has pointed out that the catalytic velocity using strong acids is generally proportional to the total concentration of the acid, and he finds this to be an important evidence in favor of complete dissociation of strong electrolytes. In applying these points of view to Goldschmidt and Thuesen's (82) investigation on esterification of weak acids by using methyl alcohol as a solvent and picric and hydrochloric acids as catalysts he finds reasonable values for the ion activity in alcoholic solutions by assuming the

hydron to be the true catalyst. Goldschmidt himself later on (80) expressed doubt as to the applicability of the dual theory in the catalysis by strong acids in alcohol.

Similar results will be obtained by going through the work of Snethlage, which has often been considered as being of particular importance for the dual theory for alcoholic solutions. Mention will be made here of his application of Bredig's diazo acetic ester catalysis in absolute alcohol at constant concentrations of picric acid on addition of varying quantities of p-toluenepicrate and related organic picrates. The velocity of reaction proved to decrease with increasing concentration of picrate as the incomplete dissociation of picric acid in alcohol was further depressed. The decrease in velocity was comparatively slight, however, and could not be made to approach zero asymptotically even at high concentrations of picric acid as ought to have been the case if the classical laws were valid and the hydron alone catalyzed the reaction. On the contrary by plotting the velocity against the reciprocal value of the picrate concentration as abscissa, Snethlage could extrapolate a value of the velocity equal to 0.013 at zero abscissa, while the acid without any picrate gave 0.058. The velocity could consequently not be reduced below 22 per cent of the initial value. The limiting value found in this way was attributed to the undissociated picric acid.

It will be shown later, however, that the dissociation of a weak acid is the more dependent upon the total salt concentration of the solution, the lower the dielectric constant of the solvent. In aqueous solution a weak acid at 0.1N salt concentration will possess a degree of dissociation twice as high as would have been the case if no salt effect had been present. The effect will be much greater when ethyl alcohol is employed as a solvent, the dielectric constant being only about one third of that for water. In this case the dissociation will be about ten times as great as calculated from the classical law of mass action. The catalytic effect at high concentrations of picrate need not therefore be due to the undissociated acid but might as well be interpreted as a hydron effect. The extrapolation to an infinitely large salt concentration is under these circumstances misleading and fails to give any information whatsoever as to the effect of the undissociated acid.

A new calculation of the experiments of Snethlage is therefore needed. Our knowledge of the dependence of the dissociation upon salt concentration in ethyl alcohol is not sufficient for such a calculation, except in extremely dilute solution, and we can therefore only indicate that these experiments are not in conflict with the assumption that the diazo ethyl acetate reaction in alcohol solutions is catalyzed only by hydriions.

The principles for calculating the shift in equilibrium, which have been of kinetic influence in Snethlage's experiments, will be given in the chapter on the secondary kinetic salt effect. It will suffice to refer to the general existence of this effect without going into further details, since such a secondary salt effect has been present in all catalytic experiments on incompletely dissociated acids in alcoholic solution, also in the exact and exhaustive material produced by Goldschmidt and his co-workers

The ideas characteristic of the dual theory for acid catalysis have been transferred to basic catalysis by S. F. Acree (4). He concludes from previous measurements in aqueous solutions and from his own measurements in alcoholic solutions that the catalytic effect is not only due to the hydroxyl ion or to the $C_2H_5O^-$ —ion but also to the "undissociated bases" like NaOH and $NaOC_2H_5$ just as in the reaction between ethyl iodide and potassium ethylate both the latter as well as the ethylate ion are supposed to take part in the reaction (63). These views have brought forth extensive and important experimental material by Acree and his co-workers, but what has been said before of the experimental data on acid catalysis as a support of the dual theory applies equally well to Acree's measurements on basic catalysis. These experiments can only be made to fit the dual theory by neglecting the secondary and primary salt effects, known to exist.

Finally the views recently given by Hantzsch (90, 91) on the nature of acid and basic catalysis should be mentioned briefly. Hantzsch criticizes severely the importance usually attached to the dissociation as a measure of the strength of the acids, although he is not really opposed to the electrolytic dissociation theory. The dissociation, which a strong acid undergoes on solution in water involves, according to his view, both a depression of the

acid properties and a levelling down of the great differences, displayed by the acids when in pure state. Hence the pure undissociated acids ought to possess a stronger catalytic effect than the dissociated aqueous solutions. This view is thus related to the dual theory of catalysis, although it was developed on a different basis.

However, while the dual theory has been restricted to the comparative study of catalysis in a definite solvent, and consequently has been able to work with the usual unprecise definition of acid and base strength, the Hantzsch theory is designed to apply to all solvents. It is quite crucial to this theory, therefore, that Hantzsch employs a completely undefined conception of acid strength. H. v. Halban (88, 89) has subjected this theory to a thorough criticism, to which here only reference can be made.

III. THE IONIZATION THEORY OF CATALYSIS

It is in many respects a reasonable idea to try to correlate the catalytic effect and the tendency of the catalyzed molecules to form more or less definite compounds with the catalyst. Such conceptions have appeared repeatedly in the development of reaction kinetics. Kastle (108) thus explains the catalytic effect of the hydron by assuming the formation of an intermediate product between hydron and catalyzed molecule and the subsequent spontaneous decomposition of the product formed. This view has been taken up by many authors, especially by Euler and Stieglitz, who have shown how to bring the theory into agreement with the experimental data.

In most of the reactions catalyzed by acids and bases, the substrate is electrically neutral and will therefore acquire an electric charge on addition of or reaction with hydrogen or hydroxyl ions. Since the assumption of such ion formation as a preliminary of the reaction itself is characteristic of the theory dealt with in this chapter this can appropriately be termed the ionization theory of catalysis. It seems practical to maintain this term also for the case of the substrate itself being already electrically charged.

As early as 1899 Euler (51, 54) launched the view that catalytic

effect is always due to an increase in the concentration of reacting ions on addition of catalyst, basing this idea upon the fact that ionic reactions as distinguished from reactions between neutral molecules proceed instantaneously or in an immeasurably short time. Related theories have also been supported by other authors (131, 117, 120). The salt effect is explained by Euler (51) as an increased ionization of the reacting substances due to a collateral increase in the dielectric constant of the system. The ideas brought forth by these authors have been important for the development of the ionization hypothesis—as demonstrated through subsequent papers by Euler and his co-workers—although a general validity cannot be given to the old conception of ionic reactivity as quoted above.

As an example of the application of the ionization hypothesis, the mutarotation of glucose may be taken, a reaction which has been subjected to a thorough analysis from different angles by numerous investigators. This reaction is catalyzed by hydrions as well as by hydroxyl ions.

The velocity at 25°C according to Hudson (99) is given by the equation:

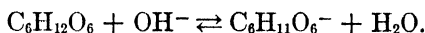
$$k = 0.0096 + 0.258 c_{H^+} + 9750 c_{OH^-}$$

where k is the monomolecular reaction constant, the time being expressed in minutes, and c_{H^+} and c_{OH^-} denote the concentration of hydrions and hydroxyl ions in the glucose solution. This equation shows that the process is composed of three partial reactions, independent of each other, viz. a spontaneous reaction independent of the acidity of the solution and two reactions catalyzed by hydrions and hydroxyl ions respectively. Riiber (155) has shown that the glucose transformation is a balanced reaction giving about 36.1 per cent α glucose and 63.9 per cent β glucose at equilibrium. In the preceding equation k is consequently equal to the sum of the two constants k_α and k_β corresponding to the two transformations $\alpha \rightarrow \beta$ and $\beta \rightarrow \alpha$ for which reason k must be independent of whether the reaction is started with the α or with the β form, as also shown by Riiber.

Euler (55, 56, 60) explains the partial reactions sensitive to

hydrions and hydroxyl ions on the basis of the ionization theory. Since glucose is able to act both as an acid and as a base, it will be ionized amphotERICALLY to glucose cations and glucose anions on addition of acid and base respectively, and these ions will in turn spontaneously break down. Other reactions catalyzed by acids or bases such as the decomposition of aceto acetic ester are interpreted by Euler (59) in the same manner.

The acid constant of glucose is approximately known, various authors (139, 116) having found values ranging between 10^{-12} and 10^{-13} . The addition of base will bring about an ionization of the glucose molecule according to the scheme:

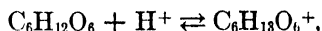


Application of the mass action law yields:

$$\frac{c_{\text{C}_6\text{H}_{11}\text{O}_6^-} c_{\text{OH}^-}}{c_{\text{C}_6\text{H}_{12}\text{O}_6}} = K,$$

showing that the spontaneous decomposition of the glucose anion obeys the same laws as a bimolecular reaction between glucose molecules and hydroxyl ions.

Basic properties may be attributed to the glucose molecule by assuming the existence of the following equilibrium:



showing in the same way that a spontaneous decomposition of the positive glucose ion yields the same course of reaction as would have been given by the bimolecular reaction between glucose and hydrions. Kinetic measurements cannot therefore furnish any evidence as to the validity of the ionization theory.

On the other hand it is not possible merely on the basis of the reaction type to refute the theory. Several facts seem to indicate that in some cases acid and basic catalysis are best interpreted from the hypothesis of ionization. It will in other cases, however, prove simpler to avoid such an explanation of the catalysis. To accept the theory as a *general* interpretation cannot be justified, the reasons being given in a subsequent chapter dealing with

a general theory of acid and basic catalysis. An essential feature in Euler's view of catalysis is his idea of ascribing approximately the same velocity of decomposition to the two ions—the glucose anion and glucose cation—reacting in acid and basic solution respectively. From that point of view the difference in the behavior of the hydron and hydroxyl ion as catalysts should be due entirely to differences in the magnitude of the acidic and basic constant of glucose.

Such an identity of the two velocities of decomposition seems *a priori* highly improbable. An experimental decision of this question is possible from the point of view of the ionization theory, provided the acid constant and basic constant of glucose are known. It would then be possible to calculate the quantities of glucose cation and glucose anion present under different conditions and then to make sure whether the ratio between these two velocities equals the velocity ratio. The basic constant, however, has not as yet been accessible to direct experimental measurement. Euler, therefore, simply assuming the correctness of this identity, calculates the ratio between the glucose anions and cations in basic and acid solutions of glucose, without of course getting any verification of his assumption by this procedure. The calculations are as follows:

The law of mass action gives for the two equilibria:

$$\frac{c_{\text{GI}^-} c_{\text{H}^+}}{c_{\text{GI}}} = K_A, \quad \frac{c_{\text{GI}^+} c_{\text{OH}^-}}{c_{\text{GI}}} = K_B$$

If the two solutions have the same c_{GI} :

$$\frac{K_A}{K_B} = \frac{c_{\text{GI}^-} c_{\text{H}^+}}{c_{\text{GI}^+} c_{\text{OH}^-}}$$

The basic constant of dissociation K_B can consequently be evaluated when K_A , c_{H^+} and c_{OH^-} are known and the ratio $\frac{c_{\text{GI}^-}}{c_{\text{GI}^+}}$ is put equal to the ratio between the velocities of the two catalyzed reactions.

On the basis of the theory of the velocity minimum the calcu-

lations are as follows. If the ionization theory is valid we may write:

$$h = k_0 c_{GI} + k_1 c_{GI^-} + k_2 c_{GI^+}.$$

The velocity minimum is determined by the equation:

$$k_1 c_{GI^-} = k_2 c_{GI^+},$$

since the product $k_1 c_{GI^-} \cdot k_2 \cdot c_{GI^+}$ is constant. By introducing K_A and K_B :

$$k_1 \frac{K_A}{c_{H^+}} = k_2 \frac{K_B}{c_{OH^-}}$$

The hydrion concentration at the minimum of velocity is consequently given by:

$$c_{H^+}^2 = \frac{k_1 K_A}{k_2 K_B} \cdot K_{H_2O}.$$

The square of the hydrion concentration at the iso-electric point is:

$$c_{H^+}^2 = \frac{K_A}{K_B} \cdot K_{H_2O}.$$

The catalytic minimum and the iso-electric point will consequently coincide, when $k_1 = k_2$, i.e. when the anion and the cation of the glucose are decomposed at the same velocity, as assumed by Euler.

The basic constant K_B is thus given by the equation:

$$K_B = \frac{K_A \cdot K_{H_2O}}{c_{H^+}^2 \text{ (minimum)}} \quad (1)$$

The point of minimum velocity, or the stability maximum, as it also has been expressed, plays an important rôle in the investigations on the glucose transformation and other reactions carried out by Euler and his co-workers. It has to be remembered, however, that the minimum is far from being sharp, the equation of Hudson corresponding to a very flat curve; secondary factors like salt effect, etc. are therefore likely to cause great displacements, while at the same time the calculations shown above are

subject to the twofold uncertainty of the ionization theory itself and of the assumption of the same transformation velocity for cation and anion. It is therefore very doubtful whether any real significance can be attached to the calculated K_B values, and equation (1) must for the same reasons be considered as exceedingly hypothetical.

Kuhn and Jacob (116) have subscribed to Euler's opinion in as far as they assume the amphoteric character of glucose and a spontaneous decomposition of the glucose ions formed upon action with the acid and with the base. But contrary to Euler, they do not assume the same reaction constant for the glucose anion as for the glucose cation, ascribing to the former about double the reaction velocity of the latter. Much stress is laid by the authors on this result which in a subsequent article (115) is made the basis of further theoretical considerations. It is therefore necessary to examine this point more closely in order to show the decisive misconception involved in the development of their results.

Both Euler, and Kuhn and Jacob present the dependence of the velocity upon the hydron and the hydroxyl ion concentration by means of a $k - p_{H^+}$ -diagram, i.e., they have employed Sørensen hydron scale as abscissa and the velocity constant as ordinate. If the catalytic effects of the hydron and the hydroxyl ion are proportional to their concentrations—and only in this case can the ionization theory be maintained—the classical theory of dissociation will demand complete symmetry of the velocity curve in the diagram mentioned. Kuhn and Jacob now allege from their measurements that the basic branch of the curve has double the slope of that of the acid part and conclude from this that the ratio between the velocities is also 2 to 1. This conclusion, however, depends upon a mathematical mistake. If Kuhn and Jacob were right in their picture of the slopes, the correct inference would be that the velocity for the basic catalysis is proportional to the square of the hydroxyl ion concentration when a first power law is followed for the acid catalysis. The curve does not say anything, however, of the relative velocities of the two catalytic reactions. The conclusions drawn by Kuhn and Jacob are therefore not only unjustified but they are also in

obvious conflict with their own equation for the catalytic process wherein proportionality with respect to both hydrion and hydroxyl ion concentration is expressed.

It is therefore not an exaggeration when Euler, Ölander and Rudberg in their studies on mutarotation (60) characterize the method of Kuhn and Jacob's calculation as "theoretisch nicht einwandsfrei." The mathematical and graphical procedure employed by these authors (60), however, does not seem very convenient either. They use a method of plotting p_{H^+} against $\log k$, in which k is the constant directly observed comprising all the three partial reactions and they seem to interpret this curve as if the two branches represented the two partial catalytic reactions. However, a simple kinetic law cannot be obtained by introducing the logarithm of a velocity constant, representing the sum of two partial constants. Such a law can be derived only by introducing the logarithm of the partial constant, i.e. by using the term $\log (k - k_0)$. On plotting this term against p_{H^+} the graph will consist of two straight lines of unit slope, showing the velocities to be proportional to the concentration of the catalysts. The curvature, displayed by the curves of Euler, Ölander and Rudberg at low concentrations of hydrions and hydroxyl ions will then of course disappear.

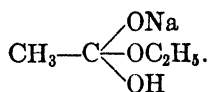
The results recorded so far are all based upon the application of the classical laws for dissociation and equilibrium, in which the gas laws are assumed to be valid for dissolved ions, thus justifying the expression of the mass action law in terms of concentrations. The data, produced hitherto on the kinetics of the glucose transformation are not sufficiently accurate for the purpose of applying the modern concepts to this reaction. An accurate experimental investigation of the reaction must, however, necessarily be accompanied by a corresponding theoretical treatment, based upon modern kinetic theories.

In this connection attention should be drawn to the fact that the great advantages of the p_{H^+} -scale in electrometric measurements, in which activity and not concentration values of the hydrogen ion are furnished are generally not present to the same extent in the field of kinetics. The tendency shown in many

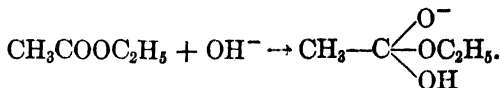
recent papers on kinetic questions to use the activity scale instead of the hydron concentration scale has only hampered the interpretation of these phenomena. The question of activity versus concentration will however be taken up in detail in a subsequent chapter, in which the point just mentioned will be further elucidated.

The same views brought forth in the interpretation of the glucose transformation have been applied by Euler and his co-workers (57, 58, 107) to other catalytic reactions especially the classical ester catalysis and the inversion of cane sugar (56). The investigation of the first reaction, which is catalyzed by hydrons as well as by hydroxyl ions, is chiefly directed towards the determination of the properties at the stability maximum.

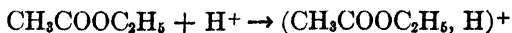
The hydrolysis of ethyl acetate in aqueous solution in presence of a base is, according to Euler's theory of ionization, dependent upon the ionization of the ethyl acetate molecule. Euler and Laurin (57) assume sodium hydroxide to react with ethyl acetate forming



This compound is presumably taken to be dissociated, so that the process really means an addition of hydroxyl ions according to the scheme:

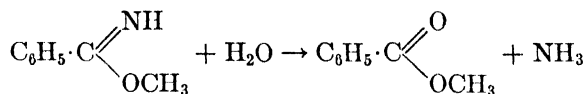


The nature of the positive ethyl acetate ion is not mentioned but can most simply be considered as a direct addition product as follows:

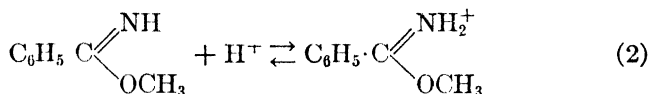


By using the mass action law it becomes evident, however, that a bimolecular reaction between the catalyst and the ester molecule and a spontaneous reaction of the positive or negative ester ion

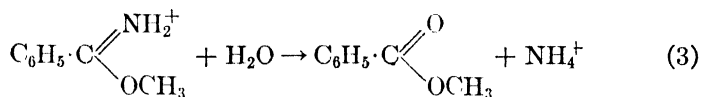
lead up to results which are kinetically identical. As in the cases previously considered an agreement between the theory and the observed course of reaction cannot be taken as a support of the ionization theory. It has been mentioned already that Stieglitz has been one of the foremost advocates of what we have called the ionization theory. He has worked chiefly on imido esters and related compounds (166). The decomposition of methylimido benzoate by water according to the scheme:



proceeds very slowly without catalysts, but is greatly accelerated in presence of acid. Stieglitz assumes the catalytic effect to be due to the formation of a positive ion:



which then reacts spontaneously with the solvent:



The catalytic effect of the hydron is consequently interpreted in accordance with the ionization theory, and a closely related explanation has also been applied by Stieglitz to the ordinary ester hydrolysis.

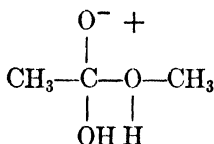
Although the kinetic basis of these views is subject to the same objections as brought out previously, Stieglitz contends that these studies constitute a crucial proof of the validity of the theory and other authors particularly Acree (2, 5, 7) have in effect supported this view. Stieglitz admits in a later paper (167), however, that the form of the reaction curve fails to give any information concerning the real mechanism of the reaction, and now resorts to other propositions for confirming his theory.

Particularly he bases his ideas upon the rule that the catalytic

effect of the hydron is closely connected with the conversion of the cation of the weaker base into the cation of the stronger base to give the highest possible diminution in free energy: Reaction (3) ought to proceed easily since NH_3 is a stronger base than

$\text{C}_6\text{H}_5 \cdot \text{C} \begin{array}{l} \nearrow \text{NH} \\ \searrow \text{OCH}_3 \end{array}$. It is easy to see, however, that Stieglitz's application of the principle of change in free energy is not correct and that his rule cannot be of general validity.

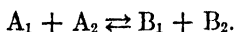
Similar considerations on the hydrolysis of esters have been brought forward by Lowry (134) who particularly endeavors to elucidate the chemical structure of the intermediate products formed by the reacting molecules. Since these intermediate compounds are supposed to react spontaneously, Lowry's views as to the mechanism of catalysis are practically congruent to those of the ionization theory. The following formula is ascribed to the active complex in the hydrolysis of ethyl acetate:



which is an ampho-ion formed simply by addition of water to the ester molecule. Lowry's concept thus seems unable to convey any immediate understanding of the catalytic phenomenon.

It has now been shown for a number of reactions that a study of the kinetic course of reaction cannot provide any support in favor of the ionization theory. This has been pointed out on the basis of the classical expressions for the velocity of reaction and the mass action law, and the result is not therefore a priori applicable to such reactions, for which the classical formulas break down. It can be shown, however, that the new and more exact theories on velocity and equilibrium in dilute solutions lead to the same result.

Consider the following chemical state of equilibrium:



The thermodynamic mass action law yields for the process:

$$\frac{a_{B_1} a_{B_2}}{a_{A_1} a_{A_2}} = K, \quad (4)$$

where a denotes the activity and K is the thermodynamic mass action constant, independent of the concentration. In the case of the A-system undergoing a change into a new system C, the velocity according to the new velocity theory is given by the formula:

$$h_A = k_A a_{A_1} a_{A_2} \cdot \frac{1}{f_X}$$

as will be explained in chapter 4.

If not A but the B system reacts, the analogous expression is:

$$h_B = k_B a_{B_1} a_{B_2} \cdot \frac{1}{f_Y}$$

where k_A and k_B are constants for dilute solutions and for the same solvent, while X and Y are the critical complexes, formed from the molecules of the A-system and the B-system, respectively. Combination of the two equations yields:

$$\frac{h_A}{h_B} = \frac{k_A}{k_B} \cdot \frac{1}{K} \cdot \frac{f_Y}{f_X}$$

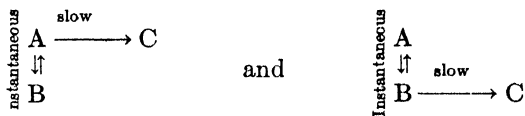
The electrical charge on the critical complexes X and Y must be equal to the total charge of the A and B systems respectively, since these complexes are formed by collision within the A- and B-systems. The total charge however, is the same, in both systems, and X and Y must consequently also carry the same charge. This means that the corresponding activity coefficients f_X and f_Y vary with the concentration in the same way as long as the solution is moderately dilute. h_A in the above equation is consequently proportional to h_B , i.e. it is not possible from the kinetic course of reaction to distinguish between the decomposition of the A- and B-systems even by using the exact equations. At higher concentrations (change in medium, addition of salt etc.) the varia-

tions in f_x and f_y will probably no longer be proportional, but this uncertainty as to the ratio between the two activity coefficients cannot of course offer any means for conclusions on the mechanism of reaction.

The above results are conclusive only if the changes $A \rightarrow C$ and $B \rightarrow C$ do not cause any disturbance in the equilibria, since equation 4 is valid only for the state of equilibrium. In other words the speed of these changes ought to be low as compared to that of the maintenance of the equilibrium $A \rightleftharpoons B$. It is of course always possible to decide whether the A- or B-system reacts, if only the reaction proceeds sufficiently slowly.

These considerations can be formulated into a general rule since they are quite independent of the particular form of the equilibrium: In any reacting system in which an equilibrium $A \rightleftharpoons B$ is instantaneously established it is impossible to decide from the course of the reaction whether the system A or the reciprocal system B reacts.

The two possibilities:



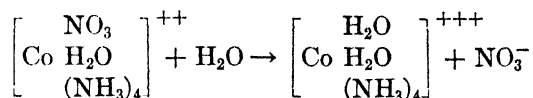
are equivalent in this respect. This proves that one must relinquish as a general proposition the idea that a verification of the ionization theory can be obtained from the kinetics of the reaction.

On the other hand it is a question of considerable theoretical interest to decide whether this equivalence between the two schemes of reaction really means that f_x and f_y are quite identical when the equilibrium is established instantaneously but this question is certainly not easy to answer.

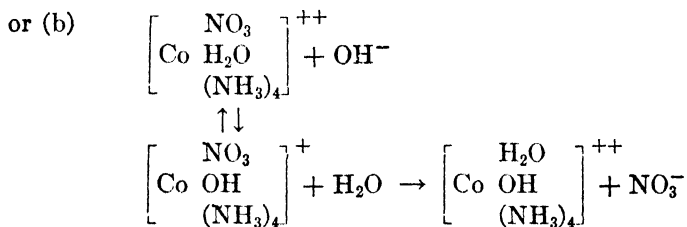
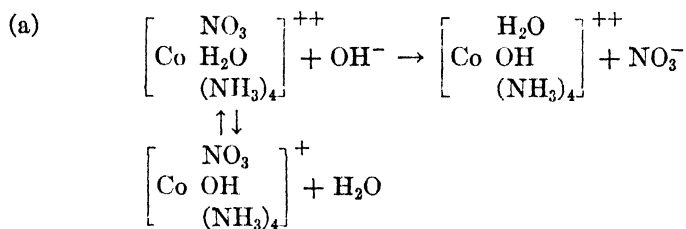
It might be possible to distinguish between the two reciprocal systems, however, when other considerations besides the purely kinetic picture so far considered are taken into account. It has been mentioned above that Stieglitz found corroboration for the ionization theory in a principle involving the connection between the strength of the base and the velocity. No general validity

could be assigned to this principle, but it might nevertheless be possible in certain cases to obtain some clues from similar considerations.

The following example from the field of the group of metal amines may be illustrative in this respect. The nitrato aquo tetramine ion is decomposed by water, splitting off the nitrato group and taking up a water molecule as follows:



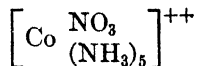
This "aquation" process is very sensitive to hydrions, the velocity in dilute acid being about inversely proportional to c_{H^+} . It might therefore seem possible to regard the reaction as an hydroxyl ion catalysis. On the other hand the nitrato ion is of slightly acid character, the hydrogen atoms of the constitutional water groups being ionizable, so the effect of diminished hydrion concentration may be to increase the acid dissociation forming the nitrato hydroxo ion which in its turn reacts spontaneously with the solvent. These possibilities may be written as follows:



where the process taken vertically represents the acid-base equilibrium which is instantaneously established. According to what

was pointed out before it is impossible to decide from the kinetic course of reaction whether reaction (a) or (b) controls the catalysis.

The fact, however, that the nitrate pentamminecobalt ion:



which is quite analogous to the nitrate aquo tetramine ion—but without acid character—shows no sensitiveness of this kind to hydrons, may be taken as strong indication that the acidity is the determining factor and, therefore the kinetic scheme is given by (b) and not by (a).

Similar deductions will surely give valuable information in many other cases as well.

More decisive criteria for the elucidation of these questions may be obtained however on the basis of new conceptions of acid and basic catalysis to be developed in the last chapter.

IV. PRIMARY SALT EFFECT. ACTIVITY AND CONCENTRATION

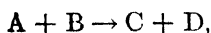
The fact that the addition of neutral salt with or without an ion in common with the catalyzing acid often causes considerable variations in the reaction constant, was one of the earliest objections against Arrhenius' hydron theory of acid catalysis. The deviations from proportionality to the concentration of a strong catalyzing acid as exhibited for instance in the inversion of cane sugar have also been considered of the same nature as the salt effect. These points have been dealt with to a certain extent in the chapter on the dual theory. In the present chapter we will discuss the salt effects more generally from the standpoint of the modern theories of electrolyte activity and of velocity of ionic reactions. It may at once be stated that the influence of salt upon the reaction velocity is far from being a kinetic abnormality even in solutions of low concentration. It is entirely reasonable to expect a kinetic effect since as already mentioned the addition of salt changes considerably the properties of water as a solvent, the activity coefficient of dissolved non-electrolytes being increased 30 to 40 per cent (157) in some cases by simply adding NaCl up to 1N. There is no reason to consider this salt effect as an

enigmatic phenomenon in conflict with the general kinetic views as seems to be the idea in Nernst's treatment (146) of this subject.

Nevertheless it is a question of great interest how this salt effect is to be properly interpreted. In this respect it is very conspicuous that the kinetic salt effects in dilute solution fall into two sharply distinguishable groups. For some reactions the effect upon the velocity constant is proportional to the concentration of the salt, in which case the effect is said to be *linear*. In other cases the effect is relatively greater the more dilute the salt solution: the variation in velocity seems rather to be proportional to the square root or to the cube root of the salt concentration, and the effect is then appropriately termed exponential in contradistinction to the linear effect (33).

A closer examination reveals that the exponential salt effect is always involved in reactions between two charged molecules, while the linear salt effect is peculiar to reactions between neutral molecules or between a neutral molecule and an ion. This observation has played an important part in the development of the general kinetic theory, in which not only the concentrations of the reacting substances but also their activity coefficients are supposed to govern the course of reaction.

It will be desirable to give a brief exposition of the general velocity theory for bimolecular reactions, since—as mentioned in the introduction—catalytic reactions can always be considered as bimolecular reactions between the molecules of the catalyst and the substrate. The velocity of a bimolecular process:



proceeding in dilute solution in an unchanging medium according to the classical theory is determined by the expression:

$$h = k \cdot c_A \cdot c_B,$$

the concentration of the reciprocal system thus having no effect upon the velocity with which A and B react. From the standpoint of velocity being related to free energy this lack of dependence upon the chemical potential of the reciprocal system may appear striking. Realizing, however, that the reaction, stoichiometrically

represented by: $A + B \rightarrow C + D$ is by no means expressed kinetically by this scheme, but actually consists in the formation of a "critical" collision complex (A, B) from collisions between A and B and a subsequent instantaneous decomposition of this complex to yield $C + D$, it appears quite natural that the chemical potential of the final system is without influence on the course of the reaction.

It is therefore essential to focus one's attention upon the process: $A + B \rightarrow (X)$, i.e. the process leading to the formation of the critical complex X . For this reaction it seems reasonable to attach great significance to the chemical potentials of all molecules taking part in the process. The chemical potential is intimately related to the activity coefficient, and the factor by which the concentration product in the classical velocity equation must be multiplied in order to take into account the influence of the activity coefficients upon the velocity, can be expressed as follows:

$$F = \frac{f_A f_B}{f_X}$$

which ratio has been called the *kinetic activity factor*.

A detailed treatment of this formula may be found in the original papers referred to previously (28, 31). It will suffice to mention here that the following hypothesis gives the foundation of the theory: "The probability of a molecular system passing from its normal state, into a state characterized by great improbability, is proportional to the ratio between the activity coefficients in the normal and the improbable state." From this rule the above kinetic activity factor is easily calculated, since the velocity of the process $A + B \rightarrow (X)$ is proportional to the probability of A and B getting into the critical state (X) . This process being measurably slow the critical state must be characterized by a high degree of improbability.

The complete expression for the velocity of the bimolecular reaction according to this theory is consequently:

$$h = k c_A c_B \frac{f_A f_B}{f_X} \quad (2)$$

For other methods leading to related expressions see Bjerrum (18) and Christiansen (39).

Expression (2) has been verified most conspicuously in ionic reactions. From what is known of the effect of ionic charge upon the activity coefficient it may be concluded that a positive salt effect, i.e., a rise in the velocity on addition of salts, is to be expected from equation (2) if the ionic reactants carry charges of the same sign, while a negative salt effect or decrease in velocity is predicted in the case of ions of opposite sign. This is exactly what is borne out by experience. The magnitude of the concentration effect in dilute solutions may be calculated from the expression of Debye-Huckel for the dependence of the activity coefficient upon the concentration (47). This theory yields the formula (36):

$$\log_{10} f = -0.5 z^2 \sqrt{\mu} \quad (3)$$

where z is the valence of the ion and μ is the ionic strength, defined for a simple salt solution by:

$$\mu = c \cdot \frac{z_1 + z_2}{2}$$

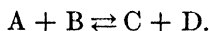
c denotes the equivalent concentration and z_1 and z_2 are the valences of the two ions of the salt. By substituting these expressions in (2) we get:

$$h = k c_A c_B 10^{z_A z_B \sqrt{\mu}} \quad (4)$$

The initial slope of the velocity-concentration curve is consequently infinite. If the equivalent salt concentration is changed from zero to 0.1 the change in velocity for two univalent ions will be about 100 per cent, and for two trivalent ions about 10^4 times the original value. These effects show what extraordinary significance must be assigned to the kinetic activity factor and demonstrates clearly the failure of the classical expressions—where the factor F is omitted—to give an even approximately correct formula for the velocity of ionic reactions. From equation (4) it can be seen that the above

mentioned law relating the type of the reaction and the sign of the salt effect must be valid, and that the *numerical* value of the salt effect is independent of the sign of the ionic charges. The formula fails, however, to give exact quantitative results except for extremely dilute solutions.

In ordinary catalysis, however, the linear salt effect is of greater importance than the exponential. Before taking up the application of the formula in the linear salt effect, the relation between the new theory and the thermodynamic mass action law as well as the part played by the activity concept in this theory should be briefly considered. First of all the formula given above fulfills the obligation of yielding the thermodynamic mass action law when applied to a chemical equilibrium, e.g.



For the velocity from left to right we have the expression

$$h_1 = k_1 c_A c_B \frac{f_A f_B}{f_X}$$

and for the reverse process:

$$h_2 = k_2 c_C c_D \frac{f_C f_D}{f_Y}.$$

At equilibrium where the velocities must be equal, we get:

$$\frac{a_C a_D}{a_A a_B} = \frac{c_C c_D f_C f_D}{c_A c_B f_A f_B} = \frac{k_1 f_Y}{k_2 f_X}.$$

According to the theory the critical complexes X and Y are, however identical; the right hand side of the equation is therefore a constant and the equation conforms with the thermodynamic law of mass action.

Another theory—the so-called activity rate theory—which also obeys this obligatory condition has been developed by H. Harned (93, 95, 96), W. C. McC. Lewis (103, 143) and other authors (158, 64). Here the velocity is assumed proportional

to the activity of the reactants, and hence for the bimolecular process $A + B \rightarrow C + D$ we get:

$$h = k a_A a_B = k c_A c_B f_A f_B, \quad (5)$$

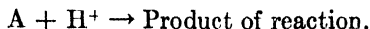
an expression which differs from (4) through the absence of the factor $\frac{1}{f_X}$.

The two theories lead to practically identical formulas, in the particular case of the critical complex being uncharged, i.e., if the two reacting ions carry charges of the same magnitude but of opposite sign. Such reactions can therefore naturally be mistaken as supporting the activity rate theory. In the case of the rearrangement of acetyl chloro amino benzene into p-chloro acetanilide, investigated by Harned and Seltz (96), the reaction may be considered to take place between the organic molecule, the H^+ ion, and the Cl^- ion. As shown by Bray and Livingston (22, 129) the same applies to the catalytic decomposition of hydrogen peroxide in an acid bromine-bromide solution, this reaction proceeding between the H_2O_2 molecule and the two ions Br^- and H^+ . Both reactions follow equation (5) closely, since in these cases f_X is equal to unity.

For all other types of reactions the formulas of the two theories are quite different. The strong positive salt effect, displayed by reactions between ions of the same sign in dilute solutions, cannot be explained by the activity rate theory—nor can the linear effect in hydron catalysis as will be shown later. In obvious conflict with facts, a negative salt effect is according to this theory, the only possible form of salt effect in dilute solutions. The reason why the activity rate theory nevertheless has found some supporting evidence is—aside from those cases where $f_X = 1$ —that the data refer to solutions of such high concentrations as to obscure to a large extent the pure activity effects. The application of this theory to ordinary acid and basic catalysis will be dealt with below.

It is of essential importance for the salt effect in normal hydron and hydroxyl ion catalysis—as exemplified by the inversion of cane sugar and ester hydrolysis—that the molecules of the sub-

strate are electrically neutral. If this be the case the bimolecular reaction between the catalyst and the substrate is given by the scheme:



Application of formula (2) will yield for the catalytic velocity:

$$h = k c_A c_{H^+} \frac{f_A f_{H^+}}{f_X},$$

f_A , the activity coefficient of a neutral molecule, being at moderately high dilutions equal to unity, while at the same time f_{H^+} and f_X approach the same ideal value. The salt effect is consequently negligible at low concentrations, retaining for the classical expression:

$$h = k c_A c_{H^+}$$

its full validity. The value of the salt effect at somewhat higher concentrations can be determined by using the expressions for the dependence of the activity coefficient upon the salt concentration. The activity coefficient of a neutral molecule seems—as far as our available data go—to be given by the equation:

$$-\ln f_0 = \beta_0 c,$$

while univalent ions at moderately low salt concentrations obey the following formula (29, 36)

$$-\ln f_1 = \alpha \sqrt{c} + \beta_1 c,$$

where α is a universal constant.

By substituting these expressions we get:

$$-\ln \frac{f_A f_{H^+}}{f_X} = (\beta_A + \beta_{H^+} - \beta_X) c,$$

and

$$h = k c_A c_{H^+} e^{(\beta_X - \beta_A - \beta_{H^+}) c} \quad (6)$$

or if the solutions are so dilute with respect to salt that the kinetic activity factor does not deviate too much from unity:

$$h = k c_A c_{H^+} [1 - (\beta_A + \beta_{H^+} - \beta_X) c]. \quad (7)$$

This equation states that the salt effect for the type of catalytic reactions under consideration varies linearly with the salt concentration, i.e., the total electrolyte concentration inclusive of the concentration of the catalyzing strong acid.

The coefficients β_A and β_{H^+} , and probably β_X as well, are of the same order of magnitude, so that the three species of molecules, involved in the reaction, can be expected to contribute equally to the salt effect. As to the absolute magnitude of β_A the evidence available from salting-out experiments has shown that the nature of both the A-molecule and the salt are of influence, so that neither the magnitude nor even the sign of the coefficient is certain. The same applies probably to β_{H^+} and β_X as well. An examination of the data available on the β -coefficient seems, however, to indicate that the following expression

$$1 > \beta_A + \beta_{H^+} - \beta_X > -1$$

is valid for most reactions, so that the maximum positive or negative salt effect would not exceed 10 per cent in a 0.1 N salt solution, while ordinarily it is considerably less. The data on kinetic salt effect confirm this supposition quite well. The fact that salt addition generally accelerates hydrion catalysis may also be explained from this formula, since β_{H^+} has a lower value than the corresponding coefficient for most other ions and since β_A , the coefficient for a non-electrolyte, is often negative.

The laws for hydroxyl ion catalysis are quite similar to those given above for acid catalysis; but the sign of the salt effect seems frequently to be negative in basic catalysis, e.g., in ester hydrolysis (9) and in the decomposition of triacetoneamine (68, 69, 35). For the individual salt effects the ionic charge is important, the catalytic effect being independent of the positive ion, if the catalyst itself is positively charged while dependent upon the negative, and vice versa (33, 40).

This theory of the primary linear salt effect may be regarded as accounting for most of the phenomena described in this chapter, without implying, however, that all other effects are of no influence. It ought especially to be emphasized that the change in medium, brought about by addition of even moderately small quantities of salt, can affect the rate of chemical reaction noticeably. The effects of changes in the medium are probably proportional to the salt concentration, and they cannot therefore be distinguished from the effects determined by formula (7).

A linear salt effect has been found in many reactions, e.g., in the inversion of cane sugar (150), ester hydrolysis (53, 151, 169), lactone formation (97, 98) etc., many of these results having been taken as supporting the dual theory as mentioned in chapter 2. More recently, the diazo acetic ester catalysis (34) and the decomposition of triacetoneamine (35) have been investigated, showing the existence of a linear salt effect—positive in the first case and negative in the latter.

The primary salt effect in the ordinary processes catalyzed by hydrions and hydroxyl ions is thus shown to be linear—i.e., proportional to the concentration—both from theoretical considerations and from experimental evidence. The conditions will of course be altered, when the substrate is no longer a neutral molecule. The effect is then determined by formula (2) as an exponential salt effect to account for the response of the ionic activity coefficients to changes in concentration. Such reactions are of great interest, but experimental data are not yet available for proving that this kind of catalysis actually exists.

The activity rate theory described previously has been based to some extent upon catalytic investigations. As the chief result in this field it has been pointed out that the activity of the catalyst and not its concentration is the essential factor in the catalysis.

This view may possibly be traced back to Lapworth (122, 102) who attaches great catalytic significance to changes in the thermodynamic potential of the hydrion, caused for example by addition of water to an alcoholic solution. In his investigation on the catalytic decomposition of hydrogen-peroxide in presence of

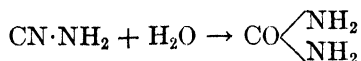
iodide ions Harned (93) claims proportionality between the rate of reaction and the chloride ion activity in solutions of sodium chloride and potassium chloride, a linear salt effect having been previously found for this reaction through work by Bredig and Walton (26). The activity of the catalyst has been introduced more directly in investigations by Jones and Lewis (103) on the inversion of cane sugar, by Harned and Seltz (96) on the transformation of acetyl chloro amino benzene into p-chloro acetanilide and by Harned and Pfanstiel (95) on the hydrolysis of ethyl acetate, these authors having emphasized the importance of the activity of the catalyst by more or less convincing arguments.

A thorough examination by Åkerlöf (1) of the acid catalysis of ethyl acetate in the presence of various salts at high concentrations has led to the following formula for the salt effect at constant hydron concentration:

$$h = C \sqrt[3]{a_{H^+}}.$$

where C is a constant. Contrary to the activity rate theory the velocity is here supposed to be proportional to the cube root of the activity and Åkerlöf's equation, applied to dilute salt solutions, cannot therefore give the catalytic activity rate theory any support.

Grube and Schmid (87) have recently subjected the neutral salt effect in the acid hydrolysis of cyanamide according to the scheme:



to a close study with a special view to strongly concentrated salt solutions. They find the simple connection:

$$h = h_0 e^{kc}, \quad (8)$$

where h and h_0 are the velocities with and without salt addition respectively, c is the salt concentration and k a constant dependent upon the nature of the salt. This equation, which shows that the same change in salt concentration produces the same

percentage change in h analogous to the equation for salting out of a non-electrolyte (157), could also be made to fit other reactions, e.g. some of the systems examined by Harned and Åkerlöf quoted previously.

It is not unlikely that an equation like (8) might have approximate validity up to quite high concentrations, in which respect the concentration scale in which c is expressed is not without consequence. It is to be noted especially that equation (8), for sufficiently small values of c , can be written in the following manner:

$$h = h_0 (1 + kc),$$

i.e., exactly analogous to the equation for the linear salt effect. The considerations, upon which the linear expression is based, lead in reality to formula (8), but the differences between the two equations can be neglected in dilute solutions. On the other hand the relation assumed by Grube and Schmid to exist between concentration and hydrion activity, and from which they conclude the importance of the latter in catalysis, is open to criticism. This relation is expressed as follows:

$$a = a_0 e^{k'c} \quad (9)$$

where a and a_0 are the hydrion activities with and without any salt addition respectively, c is the salt concentration and k' a constant dependent upon the nature of the salt. Combination of (8) and (9) yields:

$$\left(\frac{h}{h_0}\right)^{k'} = \left(\frac{a}{a_0}\right)^k$$

or

$$h = Q \cdot a^{\bar{k}}. \quad (10)$$

This formula agrees with that of Åkerlöf, if k'/k be equal to 3, which, however, according to Grube and Schmid is not the case.

Expression (9) gives at low salt concentrations:

$$a = a_0 (1 + k'c)$$

or if the concentration of acid is sufficiently low:

$$a_H = c_H (1 + k'c)$$

The hydrion activity ought therefore to vary linearly with the salt concentration—in obvious conflict with the well known connection between concentration and activity coefficient in dilute solutions. Equations (9) and (10) cannot therefore be of general validity.

Most of the experiments that have been done to construct an experimental basis for the activity rate theory of catalysis have, in order to get marked and conspicuous effects, been done at high salt concentrations. This procedure does not serve the purpose for several reasons: Firstly because the pure catalytic effect in these solutions is influenced and obscured by numerous factors, veiling the effect that is being sought. Secondly one is prevented from taking advantage of the pronounced and uniform variations with concentration shown by the activity coefficient in dilute solutions. The information obtained from a study of dilute solutions indicates beyond doubt that the important factor is the concentration of the catalyst and not its activity. The results arrived at in concentrated solutions cannot therefore be used for building up a general theory in conflict with the experimental foundation itself, but they might be combined to form empirical laws of more or less extended applicability, the experimental data being still far too scant for a satisfactory interpretation of such empirical rules.

Still one step further has been taken in support of the activity rate theory by assuming that the absolute activity of the hydrion is the governing factor in acid catalysis also when shifting from one solvent to another. This assumption in some way harmonizes with the general velocity theory developed by Dimroth (48). The fact that hydrochloric acid catalyzes lactone formation much more strongly in moist ether than in aqueous solutions of the same concentration, is explained by Taylor (172) as due to differences in the activity of HCl in the two solvents. On this point it will be sufficient to refer the reader to previous considerations (31) regarding the effect of a change in the medium.

In a conscious opposition to the concentration theory Kuhn and Jacob (116) have made the assumption that hydrion catalysis, particularly for the mutarotation of glucose, but also in general, is not proportional to the hydrion concentration but to the activity of the reacting molecules. Since their results are only of moderate accuracy and their paper does not contain any discussion at all of the mass of data contradictory to their assumption we need only quite briefly consider their statements.

Their own measurements can indeed, although with some difficulty, be interpreted in favor of the concentration theory. The difficulties lie in the authors' failure to state the acid concentration, giving only the p_{H^+} values electrometrically measured. For the most suitable experiments, viz., the measurements at $p_{H^+} = 2.09$ and 1.05 the ratio between the catalytic velocity (the total velocity minus the spontaneous) and the hydrion activity at the two concentrations is found to be 3.5 and 3.0 respectively. If the concentration is substituted for activity, the corresponding values will be 3.9 and 4.0 showing a much better agreement. The data of Kuhn and Jacob correspond, therefore, better to the concentration theory than to the activity theory.

The authors did not carry out this calculation. Their arguments for assuming proportionality with the activity are based upon their salt effect experiments. 1.0 N lithium chloride accelerates the reaction about 30 per cent in 0.1 N HCl, while the same concentration of LiCl is of no effect in 0.001N acid. This difference in the effects at different HCl concentrations is explained by the authors on the basis of changes in f_{H^+} caused by the salts. How this assumption can be made to account for the activity proportionality is not obvious.

The salt effect experiments are explained most simply by assuming the spontaneous reaction to be retarded about 10 per cent and the catalytic reaction to be accelerated about 40 per cent by 1 N LiCl. This is by no means unreasonable, since corresponding salt effects are known from analogous reactions. Both these effects are expected to be linear.

Although based upon an unsound theoretical foundation which

weakens from the outset the whole conduct of their work the paper of Kuhn and Jacob contains in some respects an interesting experimental contribution to the study of the glucose transformation as will be described further in the last chapter.

V. SECONDARY SALT EFFECT

While the primary salt effect as pointed out in the previous chapter is a direct kinetic consequence of changes in the activity coefficient of the molecules involved, the effect called the secondary salt effect (27, 38, 37, 35) is of an indirect nature, caused by a displacement in the equilibrium on addition of salt, this displacement either increasing or decreasing the concentrations of ions, participating in the reaction.

In a well known investigation on inversion velocity, Arrhenius (13) showed in 1899 that the catalytic inversion of cane sugar is much more sensitive to addition of neutral salts like KCl when the catalyst is a weak acid than when it is a strong acid. He interpreted this phenomenon as the effect of an increase in the dissociation of the weak acid and consequently also in the number of catalyzing hydrions, connecting this change in dissociation with the general deviations from the ideal osmotic laws already at that time known to exist in the case of strong electrolytes. This explanation given by Arrhenius must be considered as being correct also from the standpoint of more modern ideas. In order to get a thorough understanding of the phenomena involved, i.e., to get a quantitative theory for the secondary salt effect, it is necessary, however, to consider the case in the full light of the modern electrolytic theory.

Numerous investigators during the last decades have developed a new concept of what has been called the anomalies of strong electrolytes. Van Laar (118, 119) seems to have been the first to advance the view that salts like potassium chloride and sodium chloride are ionized in aqueous solution, to a much higher extent than concluded from conductivity data. On this basis the electrolytic conductivity cannot be taken as a measure of the degree of dissociation—as was done by Arrhenius—but the variation in the molar conductivity with concentration has to be

ascribed to the electric forces acting between the ions. The same idea has been developed independently of van Laar by Bjerrum (17), Milner (140) and Ghosh (84, 85, 86). While Milner and Ghosh have aimed at a theoretical evaluation of the effects caused by the electrostatic forces between the ions the credit for having been the first to realize the great chemical importance inherent in the theory of the complete dissociation of strong electrolytes rests with Bjerrum. He has also shown that the chemical phenomena in salt solutions are easily amenable to a simple treatment from the new viewpoints by applying the convenient concept activity—originally introduced by G. N. Lewis (126)—and the corresponding activity coefficient.

The following formula has been given by Bjerrum for the activity coefficient of an ion:

$$-\ln f = a \sqrt[3]{c} \quad (1)$$

where c is the salt concentration and a a constant dependent upon the ion charge and the type of salt. This expression, representing the logarithm of the activity coefficient as proportional to the cube root of the concentration, seems to hold quite well for concentrations from 0.01N to 0.1N, while a corresponding square root formula appears to be more exact at lower concentrations. As mentioned before the following formula (29, 36), is applicable also at somewhat higher concentrations:

$$-\ln f = \alpha \sqrt{c} + \beta c \quad (2)$$

It is beyond the scope of the present treatise to go into the details of this theory, and reference will therefore only be made to some reviews (14, 49, 100, 147) containing the chief characteristics of the theory. Concepts like "activity" and "activity coefficient" are supposed to be familiar to the reader.

An important contribution to the quantitative understanding of the conditions in solutions of strong electrolytes has been given recently by Debye and Hückel (47), who have developed theoretically an expression for the activity coefficient in dilute solutions, given in simplified form by the following formula:

$$-\log_{10} f = \frac{1.81}{(\epsilon T)^{3/2}} \cdot 10^8 z^2 \sqrt{\mu} \quad (3)$$

where μ denotes the ionic strength (see chapter 4), z the valence of the ion, D the dielectric constant of the medium and T the temperature, while at higher concentrations an additional term has to be included containing the first power of the concentration. This formula agrees well with (2), developed chiefly on an empirical basis.

The great value of these investigations in the interpretation of kinetic phenomena is due to the fact that both the kinetic activity factor and the shift of equilibrium which determines the secondary salt effect, depend upon the activity coefficients of the ions. This shifting of the equilibrium can be calculated in the following way:

If the catalyst be for example, acetic acid, dissociating according to the scheme:



the thermodynamic mass action law yields:

$$\frac{a_{\text{E}^-} a_{\text{H}^+}}{a_{\text{EH}}} = K,$$

where a is the activity. Substituting $a = cf$:

$$\frac{c_{\text{E}^-} c_{\text{H}^+}}{c_{\text{EH}}} \cdot \frac{f_{\text{E}^-} f_{\text{H}^+}}{f_{\text{EH}}} = K,$$

A change in the degree of dissociation must consequently follow when the factor $\frac{f_{\text{E}^-} f_{\text{H}^+}}{f_{\text{EH}}}$ is changed in the presence of electrolytes.

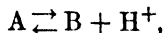
For acetic acid in 0.01N solution f_{E^-} and f_{H^+} are reduced from about 0.98 to 0.7–0.75 by addition of salt up to 0.1N, while f_{EH} may be considered as approximately unchanged and equal to unity. The dissociation of the acetic acid has consequently to increase about 40 per cent, and a similar increase has to be expected in the catalytic effect, since the hydrion concentration is the deciding factor in the catalysis.

Schreiner (161) has worked out the dissociation of acetic acid from this point of view, finding agreement with the theory, and similar verifications have been obtained in numerous other measurements (27, 28, 38).

The secondary effect exerted by salts upon catalysis by weak acids depends according to this example upon changes in the activity factor, these changes being chiefly a function of the valence type of the weak electrolyte. The influence of salt in lowering the activity coefficient is greater the higher the charge of the ion—as shown by numerous experimental data (36), and as follows from the Debye-Hückel theory. For aqueous solutions at 20°C. the Debye-Hückel formula can be written:

$$-\log_{10} f = 0.5 z^2 \sqrt{\mu} \quad (4)$$

where z is the valence of the ion. If this formula be used for calculating the activity factor in the dissociation equilibrium:



we get

$$-\log \frac{f_B f_{H^+}}{f_A} = (z_B^2 - z_A^2 + 1) 0.5 \sqrt{\mu}$$

or

$$\log \frac{f_B f_{H^+}}{f_A} = z_B \sqrt{\mu}$$

and the corresponding mass action expression becomes:

$$\frac{c_B c_{H^+}}{c_A} = K \cdot 10^{-z_B \sqrt{\mu}} \quad (5)$$

The dissociation constant is consequently independent of the salt concentration when $z = 0$, i.e., when the B-molecule is uncharged. Hence the equilibrium

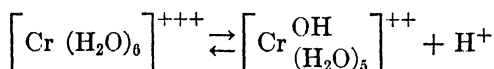


where $B = NH_3$, will not be shifted by addition of salt at moderately low concentrations. The acidity created by the above dissociation scheme or by a dissociation of the same electrical type will therefore be governed by particularly simple laws.

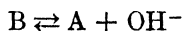
If z_B is negative, salt addition will evidently increase the dis-

sociation of the acid as shown previously in the case of acetic acid. This effect will be designated a positive salt effect. It will be still more pronounced, if the charge of the acid itself is negative, as for example for the acid oxalate and tartrate ions.

If z_B is positive, the dissociation of the acid will decrease on addition of salt, the salt effect being in this case negative; as an example one may take the "hydrolytic" decomposition of the chromic ion (38), which can be considered most simply as a dissociation according to the scheme:



The secondary salt effect in hydroxyl ion catalysis can be calculated in a similar way by considering the equilibrium:



The corresponding equations will be:

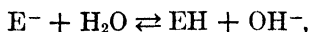
$$\frac{c_A c_{\text{OH}^-}}{c_B} \frac{f_A f_{\text{OH}^-}}{f_B} = K,$$

$$\log \frac{f_A f_{\text{OH}^-}}{f_B} = -z_A \sqrt{\mu},$$

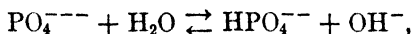
and

$$\frac{c_A c_{\text{OH}^-}}{c_B} = K \cdot 10^{z_A \sqrt{\mu}}, \quad (6)$$

A positive salt effect ought consequently to be obtained, when z_A is positive, for example when the base is uncharged like NH_3 or when it carries positive charges. The salt effect is zero when $z_A = 0$, for example in the equilibrium:



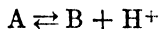
and finally negative salt effect has to be expected, when z_A is negative, e.g.:



when $z_A = -2$.

As long as the solutions are sufficiently dilute with respect to salt, for example not exceeding 0.01N, the formulas (5) and (6) will permit of a fairly exact quantitative calculation of the salt effect. The deviations appearing at higher concentrations are probably as a first approximation best given as varying linearly with the salt concentration, the proportionality factors involved being of the same order of magnitude as the salting-out coefficient for non-electrolytes and the coefficient β in formula (2).

It should be noted that the magnitude of the effect here discussed is largely a function of the ratio in which the molecules are present in the equilibria considered. Three separate cases might here be distinguished. Consider the equilibrium:



and suppose the catalysis to be due to the hydrions. If $c_B = c_{H^+}$ (I), that is the solution can be made up from the weak electrolyte A and the solvent alone the salt effect will fall upon B and H^+ to the same extent. If $c_B \gg c_{H^+}$ (II), i.e., in the case of a buffer solution, the whole change in concentration will be restricted to the hydrion, and the catalytic effect will be double that in case (I); and finally if $c_{H^+} \gg c_B$ (III), the solution is regulated with respect to B-molecules and the secondary salt effect equal to zero.

It is evident from what has been stated that the secondary salt effect is dependent upon a number of factors and so may prove of a rather complicated nature. It is important therefore that the theory has been amply verified. The law of secondary salt effect is obeyed not only when a salt like KCl is added to acetic acid in the inversion of cane sugar or for the hydrolysis of esters [Arrhenius (13, 27)] but also in cases where salts with a common ion are added to weak acids or bases. The older data show on closer examination that considerable deviations from the simple classical laws exist. For instance when sodium acetate is added to acetic acid or when ammonium chloride (10, 35) is added to ammonia results are obtained which conform to the requirements of the secondary salt effect.

Furthermore the influence of differences in the type of dis-

sociation has been verified by recent investigations. A strong negative salt effect has been found for the basic catalysis of nitroso triacetoneamine (35), when the reaction proceeds in a mixture of secondary and tertiary phosphates. A negative effect has also been found for the diazo acetic ester catalysis in solutions of chromic salts (38). These phenomena are all in agreement with the theory of the secondary salt effect.

The considerations made in this and the preceding chapter in connection with the development of the theory of primary and secondary salt effects clearly show the inadequacy of attempting to relate the salt effects simply to the p_{H^+} values of the solutions. These values are only of indirect interest in the question of the rate of reaction—as pointed out before, while at the same time they fail to take into account the general change in medium brought about by shifting the equilibria in the solution. It has been overlooked in the numerous attempts of this kind not only that the p_{H^+} -function (or potential in general) is but indirectly of significance in the question of reaction rate, but also, that the effect of added salts is, as shown, very pronouncedly dependent upon the general nature of the system under investigation.

VI. THE ACID-BASIC FUNCTION

For a complete understanding of the phenomena in acid and basic catalysis, an analysis of the very concept of acids and bases is necessary. The historical development of this concept, as outlined for example in S. M. Jørgensen's (104) and Ramberg's (152) expositions of the subject, gives an impression of very fluctuating views in the course of time. Here we are not going to deal with the historical evolution but try to present the ideas to which recent studies of the nature of acids and bases have led.

The definition of an acid deduced from Arrhenius' dissociation theory may be considered as the one generally accepted nowadays. According to this any substance splitting off hydrions on solution in water is an acid. The definition in this form is hardly consistent since certain properties are attributed to the substance

under circumstances different from those under which the observations are made. A substance ought to be pronounced an acid under certain conditions only if hydrions are released from it under the same conditions. Pure acids like acetic acid and nitric acid will perhaps not be considered as acids by this definition, if hydrions are undetectable when the substances are in a pure state, i.e., in the absence of solvent.

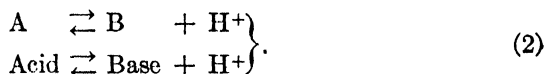
It is quite possible, however, that a substance which does not give off hydrions under certain conditions can nevertheless, possess properties which justify the application of the term acid. This is easily understood when realizing that the reaction indicating the acidity of the acid:



is not identical with the reaction governing the dissociation in a solvent, since it is not the free but only the solvated hydrions which attain perceptible concentrations in solution. The real dissociation always proceeds in cooperation with the solvent, and the degree of dissociation is dependent upon the readiness with which this cooperation is afforded. In defining an acid, however, it is preferable, to disregard this influence of the solvent and instead of the true process of dissociation to use scheme (1) as our definition of the acid A, expressing merely the splitting off of a hydrion, without saying anything concerning its subsequent fate.

Now considering the definition of a base it is well known that, since the formulation of the dissociation theory this has been closely associated with the hydroxyl ion. A substance, according to this view, is considered a base when it splits off hydroxyl ions in aqueous solutions. This definition, however, is much too narrow, even after being modified to include the solvent effect as in the definition of an acid given above. Those substances, designated as bases, are, quite generally and independently of the solvent, characterized by their *ability to take up hydrions*. Scheme (1) will consequently express both the acid character of the molecule A and the basic character of the mole-

cule B and can therefore be used as the schematic definition for both; namely,

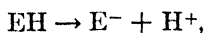


A base is necessarily formed whenever a substance functions as an acid giving off hydrions, and conversely an acid has always to appear when a base takes up a hydrion. Acids and bases related to each other by such a scheme are designated as conjugate acids and bases.

The advantages of this definition are developed elsewhere (30) and will not be treated in detail here. It suffices to mention that the scheme gives in all cases a logical and adequate expression of the properties which are generally recognized as being characteristic of acids and bases. The close relation between these two groups of substances is explained by their fundamental relation to the hydrion, while their unique position as compared with all other chemical substances is due to the uniqueness of the hydrion, consisting solely of an atomic nucleus. The only possible parallel to this class of substances is the oxidation-reduction system, in which the electron plays the same rôle as does the hydrogen nucleus or proton in the acid-base-system.

Furthermore the thermodynamic significance of the scheme (2) ought to be emphasized. Even if the two opposed reactions in the equilibrium given by (2) fail in giving the correct kinetic picture of the process, the thermodynamic mass action law may nevertheless be applied to the equilibrium, irrespective of whether or not actual existence is assigned to the hydrion. This point will be taken up in the next chapter in the problem of obtaining a suitable standard for the concepts of acidity and basicity as well as of acid and basic strength.

There can be inferred from scheme (2) as a result of particular importance that conjugate acids and bases cannot carry the same electric charge. If the acid is electrically neutral, the base must be negatively charged, e.g., in the acetic acid-acetate ion system.



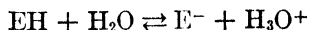
while a neutral base requires a positively charged acid, e.g.



In this respect the new definition shows a distinctive departure from the old theory in which the electric neutrality of the acid and basic molecules was considered an essential feature (137). According to the new concept molecules exert their acid or basic functions independently of their electric charges, the acid and basic properties being ascribable in the same way to ions and uncharged molecules.

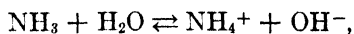
The recognition of this point of view implies an unusual extension of the acid-base concept, but at the same time creates a simplification and conformity in the treatment of the phenomena involved both in aqueous and non-aqueous solutions. Reference is made to earlier papers on this subject (30, 32).

The fact that a molecule may exhibit acidity or basicity independent of its charge does not mean, of course, that the charge is without influence upon the strength of the acid and the base. Nor does the omission of the solvent in scheme (2) imply that the solvent is immaterial for the manifestation of the acid and basic properties. On the contrary, as shown in chapter 7, the influence of the solvent is of the greatest importance. First of all the dissociation of an acid, as mentioned, takes place only under the influence of the solvent. For acetic acid in water the dissociation equilibrium will be given by the scheme:

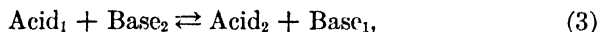


in which EH and H_3O^+ according to (2) act as acids and H_2O and E^- as bases. It is reasonable to assume that the hydrion unites with one molecule of water, just as it takes up one molecule of ammonia. This is in agreement with the results obtained by numerous authors (83, 62, 90, 63, 160, 134, 32). The normal, hydrated hydrion, or the oxonium ion, may therefore be considered as an acid, possessing special significance not in any essential respect but only because of the practical consideration that its conjugate base is the most common solvent, water.

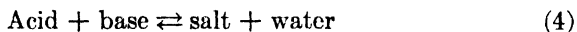
Addition of the base, ammonia, to the solvent water will in the same manner give rise to the equilibrium:



where NH_3 and OH^- function as bases and H_2O and NH_4^+ as acids. The different rôle played by the water in these two examples makes clear its amphoteric character and may serve as an example in considering the *double acid-base equilibrium*, which will be established in a system comprising two pairs of conjugate acids and bases:



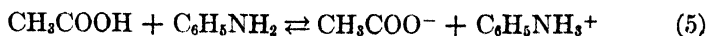
In conformity with this scheme, the reaction taking place on mixing an acid and a base consists in the formation of a new acid and a new base. This result is in marked opposition to the usual ideas according to which the resulting reaction is a "neutralization."



Since, however, the concepts of "neutralization" and "neutrality," as has been shown before (30), are illogical and misleading, it seems desirable to subject scheme (4) which is fundamental to these concepts to a closer examination especially as to its relation to (3). One would expect that such an analysis would lead to a more satisfactory understanding of the "salt" concept and its relation to that of acids and bases than has hitherto been obtained.

The extent, to which acids and bases react upon mixing, is as is well known entirely dependent upon their strength. For "strong" acids and bases in aqueous solution the reaction is practically complete; for acetic acid and ammonia it is appreciably incomplete; for acetic acid and aniline only 50 per cent complete; while finally phenol and aniline scarcely react at all. Theoretically the reactions can be considered as incomplete in all cases, and the nature of the acid and basic properties is consequently analyzed most generally by taking acids and bases which do not exhibit extreme effects.

For that purpose consider the reaction between acetic acid and aniline



and the reaction between acetic acid and the propionate ion:



The extent of reaction in both these cases is about 50%, so that the concentrations of the two reciprocal systems are about the same at equilibrium. Scheme (5) taken from left to right denotes the reaction between a neutral acid and a neutral base and in the opposite direction a reaction between an ion base and an ion acid, while scheme (6) gives in either direction the reaction between an ion base and a neutral acid. Both reactions conform completely to scheme (3).

It is immediately apparent by comparison with scheme (4) that the water molecule is absent in (5) and (6). This might perhaps be considered as immaterial, since in aqueous solution it is always possible to introduce a water molecule in the equation by assuming hydration to take place. The effect of such a water molecule could not be distinguished, however, from the ordinary solvation effect, pertaining to all reactions proceeding in a solvent, and therefore cannot furnish a distinguishing characteristic in acid and basic reactions.

Furthermore it is seen in the case of reaction (4) that acetate ion + anilinium ion, i.e., acid and base with opposite charges, serve as a substitute for the "salt" in scheme (4). The definition of a salt given by (4) corresponds consequently in this respect to the ordinary concept of salts as substances being ionized into two or more oppositely charged ions. The product of reaction in (6), however, cannot possibly be covered by the salt concept, nor does the product in scheme (5), taken from right to left, viz. acetic acid + aniline, justify the application of this term.

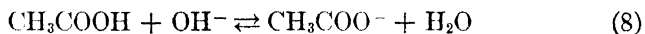
Hence "salt formation" in reactions between acids and bases becomes restricted to cases, in which the acid-base system reacting is uncharged. If it be recognized, therefore, that the

electric charge does not affect the characteristic properties of acids and bases, the conclusion seems inevitable that also "salt formation" is deprived of the rôle hitherto assigned to it in acid-basic reactions thus leaving the relation between salts, acids and bases involved in scheme (4) without general significance.

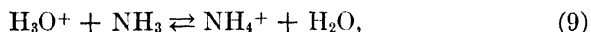
The usual scheme (4), comprising the old views on acid and basic reactions, is therefore certainly not adequate according to the new concepts. It ought to be stated, however, that this scheme has originated from considerations on "strong" acids and bases in aqueous solution. Since the base in this case is the hydroxyl ion, the "process of neutralization" will certainly be accompanied by the formation of water. This is naturally not in conflict with our point of view, according to which H_2O might take part in the reaction just as well as any other base and acid. If both the acid and the base are strong, the whole reaction will correspond to a water formation, thus:



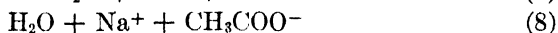
and if the acid is weak and the base strong:



or the acid strong and the base weak:



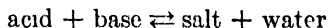
water and no salt will be formed. Previously it has been customary to consider $\text{Na}^+ + \text{OH}^-$ and not OH^- alone as the base; in the same way the term acid has been applied to $\text{H}_3\text{O}^+ + \text{Cl}^-$ and not to H_3O^+ alone. The products in the above equations would then be:



in agreement with the old idea of a salt formation. It seems self-evident, however, that an admixture of ions—already present in the original solutions—can hardly be characteristic of a

chemical process, and also that the systems $\text{H}_3\text{O}^+ + \text{Cl}^-$ and $\text{Na}^+ + \text{OH}^-$ deserve the designation of salts just as much as do $\text{Na}^+ + \text{Cl}^-$, $\text{Na}^+ + \text{CH}_3\text{COO}^-$ or $\text{NH}_4^+ + \text{Cl}^-$. The acid-base reaction would then according to the old point of view hardly represent a salt formation, but rather a salt transformation.

Accordingly, since we are justified in concluding that the usual scheme:



does not express the essential and characteristic features in the reaction between acids and bases, it becomes evident that the conception of a salt—if at all accessible for a proper delimitation—has to be established by definitions which have no relation to the conception of acids and bases.

E. Franklin (70) has pointed out that the common scheme (4) for acid, base and salt can be replaced in many reactions by other schemes in which the part played by H_2O is taken by other substances such as HF , H_2S , H_3N , etc., thus making a distinction between the "water system," the "hydrogen fluoride system" etc. for acid, bases and salts. These considerations being primarily systematic in character, do not add anything new concerning the characteristic functions of acids and bases. The objections brought against the ordinary ideas on the "water system," may be transferred unchanged in principle to Franklin's systems.

VII. ON THE STRENGTH OF ACIDS AND BASES

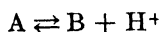
To the considerations on the acid-basic function given in the preceding chapter it is necessary to attach a definition of acid and basic strength as well as of the concepts of acidity and basicity. These terms have been confined primarily to aqueous solutions or, in any case, to solutions of constant environment, without much consideration as to how the concepts can be extended to include variations in solvent. Since in aqueous solutions acidity is most conveniently defined as the hydrion activity or the hydrion potential, it would appear reasonable to define the acidity of any homogeneous system as the hydrion potential in that system. Such a definition would be similar to the one suggested by Lap-

worth (123), whose views on this point, however, are far from being clear. Basicity is then the reciprocal value of the activity or the negative value of the hydrion potential.

Measurements of acidity and basicity, defined in this manner, could then, in principle, be carried out by determining the potential difference between two reversible hydrogen electrodes, one dipping into a standard solution of unit hydrion activity and the other into the solution in question, both solutions being brought to the same electrical potential. The activity concept to be used here is the same as the one originally proposed by G. N. Lewis, identity of activity of a substance in two solvents indicating that also its chemical potential is the same in these solvents. This activity may be called the absolute activity in order to avoid ambiguity. The activity of an acid is of course an entirely different concept from that of acidity since the former depends not only upon the H^+ ion but also upon the corresponding anion.

While fixing the conception of acidity and basicity by a theoretical definition is thus of no difficulty the problem of obtaining a general measure of *acid and basic strength* may appear a little more complicated. In conformity with the dissociation theory, it has become customary to determine the strength of an acid by its dissociation constant in dilute aqueous solution. This method is excellent also for non-aqueous solutions, as long as the problem is only to compare the strength of various acids in the same medium. Constancy of the medium obviously requires that the concentration of the acids under investigation should be very low, since a solution of finite concentration really is a medium different from the pure solvent.

The "dissociation constant" is however no longer a simple measure of acid and basic strength, when the medium is changed. It would therefore hardly be rational to use this constant for comparing acid strengths in different media. In searching for a more general measure of the strength, it would seem expedient to relate the question of strength to the general concept of acidity and basicity. Keeping in mind our scheme of definition:



we will decide to measure the strength of acids and bases by the acidity and basicity, respectively, exhibited in the solutions under comparable conditions, i.e., when the ratio between the concentrations of acid and base is a constant. If unity is chosen for this constant ratio, we get the expression:

$$\frac{c_B}{c_A} \cdot a_{H^+} = K_{\text{Acid}} \quad (1)$$

as a measure of the acid strength for variations of both the acid and the medium, and similarly for the basic strength

$$\frac{c_A}{c_B} \cdot \frac{1}{a_{H^+}} = K_{\text{Bas.}} \quad (2)$$

The constants thus defined, for which the term *acidity constant* and *basicity constant* should be used, are—as a closer study shows—well adapted as measures of acid and basic strength.

In the above equations c refers to the stoichiometric concentrations of the acid and base involved, irrespective of their solvation in the solution. It is therefore very probable that what are called c_A and c_B for the same acid-base system in different solvents actually correspond to entirely different A- and B- molecules. It might perhaps have been preferable in the definition of $K_{\text{Acid.}}$ and $K_{\text{Bas.}}$ to let c_A and c_B denote the concentrations of the real, free A- and B-molecules, if it were not for our limited knowledge of the state of solvation, which does not suffice for calculating the concentration on this basis.

The above constants differ from the usual constants of dissociation by containing a_{H^+} —the absolute activity of the hydrion—instead of the hydrion concentration. One might wonder whether the substitution for c_A and c_B of the corresponding absolute activities would carry the problem still a step further. The strength constant defined in this way:

$$\frac{a_B}{a_A} a_{H^+} = K_{\text{Act.}}$$

is a real thermodynamic constant, the activity constant, quite independent of the medium, and this K might therefore seem suitable as an absolute definition of the acid strength.

A closer study, however, will easily reveal the failure of this method. The quantities involved in K_{Acid} viz. the absolute activity of the hydron and the stoichiometric concentrations of A and B, are comparable for variations both of the acid and of the solvent, whereas the activities a_A and a_B entering in the definition of K_{Act} can quite well be compared for the same acid in different media, but by no means for different acids in the same medium. It is for instance impossible to measure the activity of hydrochloric acid in relation to the activity of acetic acid, before both of these values have been fixed by an arbitrary convention. Then, however, the strength constant K_{Act} will become a function of this convention and consequently will not afford the absolute measure sought for. For instance, this standardization might be brought about by assuming activity and stoichiometric concentration to attain the same value in ideally dilute aqueous solutions, as is usually done in the thermodynamic treatment of electrolytes dissolved in water. K_{Act} would then be identical with K_{Acid} in the special case where water is the medium. A more general measure of acid and basic strength would be possible only if a particular state were found, which could represent a sort of general and ideal reference state. The attenuated vapor is actually such an ideal state, and from a theoretical point of view the best solution of the problem would be to refer the acid and basic strength to the vapor state. The activities of A and B would then be equal to the concentrations in the vapor system, and the acid strength would be the acidity in an attenuated vapor phase at equal concentrations of A and B. A practical application of this principle is precluded, however, since at least two different species of electrically charged molecules are involved in the acid-basic equilibrium.

Maintaining therefore the measures of acid and basic strength given above, viz:

$$\frac{c_B}{c_A} \cdot a_{H^+} = K_{\text{Acid.}}, \quad \frac{c_A}{c_B} \cdot \frac{1}{a_{H^+}} = K_{\text{Bas.}}$$

the next step will be to investigate more closely the influence of medium and ionic charge type upon these constants.

By introducing the relation between the absolute activity a , the absolute activity coefficient f and the concentration:

$$a = cf,$$

the expression for the acidity constant may be transcribed as follows:

$$K_{\text{Acid.}} = \frac{c_B}{c_A} \cdot a_{H^+} = \frac{a_B}{a_A} \cdot a_{H^+} \cdot \frac{f_A}{f_B},$$

in which $\frac{a_B}{a_A} \cdot a_{H^+} = K_{\text{Act.}}$ is a thermodynamic constant for each

particular acid. Variations in the acidity constant by changing from one medium to another are consequently determined only by the ratio between the absolute, stoichiometric activity coefficients f_A and f_B and are notably not—as shown later to be the case for the dissociation constant—dependent upon the acid or basic character of the medium. The stoichiometric coefficients f_A and f_B include all the effects—especially also the solvation effect—produced by a change in medium, for which reason the variations in f_A and f_B with the medium are undoubtedly of an individual character. Nevertheless it can be stated as a general rule governing the dependence of this change upon the ionic charge, that the higher the charge on the ion the greater will be the increase in its activity coefficient on changing from a medium of high dielectric constant to one of lower dielectric constant. If a series of acids of different electric charges: A^- , A^0 , A^+ , A^{++} etc. be considered, the changes in f for these acids and the corresponding bases, when the medium is changed for example from water of high dielectric constant to alcohol of lower dielectric constant, will come out according to the following scheme:

Acid	Base	f_A	f_B	$\frac{f_A}{f_B}$	$K_{\text{Acid.}}$
A^-	B^-	+	++	—	—
A^0	B^-	0	+	—	—
A^+	B^0	+	0	+	+
A^{++}	B^+	++	+	+	+

where an increase is denoted by + and a decrease by -. This scheme represents clearly the effect that the medium exerts upon the acidity constant by virtue of its dielectric constant, since the ratio $\frac{f_A}{f_B}$, as shown previously, determines the change in K_{Acid} .

Thus on proceeding from water as a solvent to alcohol, the constant for acetic acid of type A° will decrease, while it will increase for the ammonium ion, which is of type A+. The change will of course be opposite for the basicity constant, which is the reciprocal of the acidity constant of the conjugate acid.

The above rule that a decrease in the dielectric constant produces a higher increase in the activity of the ion, the higher the numerical value of the electric charge, follows simply from the electrostatic formula for the potential of a spherical particle in a medium of dielectric constant D:

$$\psi = \frac{\epsilon z}{Dr}$$

where ϵ is the unit charge of electricity, z the number of these charges on the particle and r the radius.

The electrical energy of the particle is then:

$$E = \frac{\epsilon^2 z^2}{2 Dr}$$

and the work which can be gained by transference from medium 1 to medium 2:

$$A = \frac{z^2 \epsilon^2}{2r} \left(\frac{1}{D_1} - \frac{1}{D_2} \right). \quad (3)$$

This work may also be written:

$$A = k T \ln \frac{f_1}{f_2} \quad (4)$$

where f_1 and f_2 are the absolute activity coefficients in medium 1 and 2, respectively. Hence:

$$\ln \frac{f_1}{f_2} = \frac{z^2 \epsilon^2}{2 rk T} \left(\frac{1}{D_1} - \frac{1}{D_2} \right) \quad (5)$$

By introducing:

$$K_{\text{Acid}} = K_{\text{Act.}} \frac{f_A}{f_B} \quad (6)$$

and

$$z_A = z_B + 1$$

we get:

$$\ln \frac{K_{\text{Acid (1)}}}{K_{\text{Acid (2)}}} = \frac{\epsilon^2}{2 k T r} \left(\frac{1}{D_1} - \frac{1}{D_2} \right) (2 z_B + 1), \quad (7)$$

where $r_A = r_B = r$.

Max Born (21) has made use of equation (3) for calculating ion sizes by assuming the work performed in the transference from vacuum to water to be equal to the heat of hydration. For calculation of the change in the acidity from water to methyl alcohol numerical values are now introduced:

$$\begin{aligned} \epsilon &= 4.77 \cdot 10^{-10} \\ k &= 1.37 \cdot 10^{-16} \\ D_1 &= 81 \\ D_2 &= 31 \\ T &= 290 \\ r &= 2 \cdot 10^{-8} \end{aligned}$$

the latter value having of course only an approximate and uncertain character. By substitution:

$$\begin{aligned} \log \frac{f_1}{f_2} &= 62.2 \left(\frac{1}{D_1} - \frac{1}{D_2} \right) \\ \log \frac{f_1}{f_2} &= 1.24 z^2 \end{aligned}$$

and

$$\log \frac{K_{\text{Acid (Methyl alc.)}}}{K_{\text{Acid (Water)}}} = 1.24 (2 z_B + 1).$$

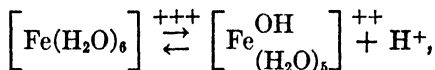
The acidity constant for acetic acid, where $z_B = -1$ and for the ammonium ion where $z_B = 0$ ought accordingly to decrease and increase, respectively, by 1.24 logarithmic units on passing

from water to methyl alcohol, i.e. the acid should become 17 times weaker or stronger by this change in solvent. The observed change agrees always in sign with the calculation, but the numerical value found by experiment is much higher than the calculated in spite of the fact that the value used for r is probably too small. There is nothing surprising in this disagreement, because the simplified picture of a molecule used in these calculations is certainly very far from representing its actual structure.

It has been pointed out previously (32) that the effect of the medium upon the dissociation constant may be estimated by considering the influence of a change in the dielectric constant upon the forces between the dissociation products of the acid, i.e. between the base and the hydrion. This effect depends upon the type of dissociation. On proceeding from a medium of a higher to a medium of a lower dielectric constant the forces between the charged particles will increase and the dissociation consequently decreases if the base is negatively charged and increases if the base is positively charged. The diminution in the dissociation of acetic acid:



brought about by changing the medium from water to aqueous alcohol, is therefore to be attributed largely to a corresponding rise in the electric attraction between the ions of the acid; conversely the increase in the dissociation of ferric ion:



caused by addition of alcohol, is explained on the basis of an increase in the repulsive forces between the positive hydroxo-ion and the hydrion. It can be shown that this mode of attacking the problem gives the same quantitative results as the previous method.

As long as only the electric effect is taken into account the dissociation constant on changing from one medium to another may be considered as inversely proportional to the number of collisions per second between the basic molecule and the hydrion. Using

the method given by J. A. Christiansen (39) the influence of the electric forces upon the number of impacts will be determined by Boltzmann's factor:

$$e^{-\frac{e}{kT} \psi_B}$$

where ψ_B , the potential at a distance r from the center of the B sphere, is given by the equation:

$$\psi_B = \frac{z_B e}{Dr},$$

r being the sum of the radii of the B-molecule and the hydrion. The influence of the electric forces upon the number of collisions is thus governed by the factor:

$$e^{-\frac{z_B e^2}{k T D r}}$$

and consequently:

$$\ln \frac{K_{\text{Dis. (1)}}}{K_{\text{Dis. (2)}}} = \frac{e^2}{kT r} \left(\frac{1}{D_1} - \frac{1}{D_2} \right) z_B \quad (8)$$

This formula differs from (7) by containing the factor $2z_B$ instead of $(2z_B + 1)$. The reason for this difference is that (7) contains the acidity constant, whereas K in (8) stands for the following constant:

$$\frac{c_B}{c_A} c_{H^+}$$

the ratio between this constant and the acidity constant being $\frac{1}{f_{H^+}}$. By introducing (5):

$$\ln \frac{f_{H^+ (1)}}{f_{H^+ (2)}} = \frac{e^2}{2rkT} \left(\frac{1}{D_1} - \frac{1}{D_2} \right)$$

(7) and (8) are seen to be identical.

These formulas for the ratio between the absolute activity coefficients in two solvents usually yield as already mentioned values

which are too low. These deviations are due first of all to the unjustified extension to real ions of the electrostatic laws for spherical particles but account ought also to be taken of both the "physical" and "chemical" solvates existing in solution in many cases.

Numerous authors have been concerned with the question of the dependence of dissociation and chemical potential upon the properties of the solvent. It is well known that Walden (177) has given the following empirical relation:

$$\frac{c_1}{c_2} = \left(\frac{D_1}{D_2}\right)^3$$

between the dielectric constants and the concentrations, for which a salt shows the same dissociation in two solvents. A corresponding equation has also been developed on theoretical grounds (16, 114). This formula gives qualitatively the same dependence of the ion potentials upon the medium as do the above equations, although the lack of concordance is otherwise obvious. Numerous experiments on solubility and E.M.F. (101, 144, 125, 94, 160) give further qualitative confirmation of these results.

Experimental data, on the influence of the medium upon what we have called the acidity and basicity constants are only meagre. The investigations have generally been concerned with the ordinary dissociation constant as for instance in recent papers by Schreiner (162) on strong acids in water, methyl and ethyl alcohol and by Bjerrum, Unmack and Zechmeister (19) on electrolytes in methyl alcohol.

A verification of the views developed here for the influence of the medium upon the acidity constant can be found, however, in a series of investigations by Michaelis and Mizutani (138, 141, 142). They determine electrometrically the hydron potential in buffer mixtures in water and in water-alcohol solutions and compute from these measurements an expression identical with the acidity constant. A rise in the alcohol percentage from 0 to about 90 causes regularly a *decrease* in the acidity constant (2-3 logarithmic units) for neutral acids (acetic acid, benzoic acid etc.) and a corresponding, although numerically somewhat smaller, *increase* in

the constant for positive ions (ammonium ion, anilinium ion etc.)—which is in good agreement with the views developed above. The changes in the dissociation of the negative acids and the amphoteric character of the amino benzoic acids are likewise in accordance with the theory. In the case of glycine the first dissociation step (the dissociation of the positively charged glycinium ion) shows a decrease in acidity on addition of alcohol, whereas the second step is not materially affected by the addition, this behavior being explained when the amphoteric form of the glycine and ordinary singly-charged ions are supposed to possess a similar sensitivity to alcohol. In agreement with this view glycine which is highly soluble in water is mentioned as insoluble in absolute alcohol.

Whether the assumption made by Michaelis and Mizutani that the acid character of the carboxyl group is more sensitive to alcohol than is the basic character of the amino group is a question which cannot yet be definitely decided.

The acidity and basicity constants are, as mentioned before, subject to more individual effects in addition to those arising from variations in the dielectric constant and the electric charge of the acid and the base. The chemical potential of a particle in a solution depends upon the surroundings through influences other than those of the electric net charge. For some non-electrolytes the potential is for example higher in water than in benzene, while the opposite may be true in other cases. Such effects must of course be operating quite generally, an increase in the absolute activity coefficient of the acid causing the acidity constant to increase, whereas the constant will be influenced in the opposite manner by a corresponding rise in the activity coefficient of the base. For non-electrolytes the individual effects are easily accessible to observation. In the case of ions electromotive force measurements—as in the papers quoted above—may be applicable but high and difficultly accessible boundary potentials will most likely obscure the results. Comparing a series of salts with a common ion, e.g. in respect to their solubilities, it is possible to obtain data for the activity ratio of the non-common ions in the two solvents. The ratio for a single ion is however not accessible by this method.

The fact that the anilinium ion and particularly the dimethylanilinium ion shows greater increase in the acidity constant from water to alcohol than does the electrically analogous ammonium ion—as found by Michaelis and Mizutani—represents an individual activity effect which becomes intelligible in view of the considerable increase in the solubility ratio amine/ammonia brought about by changing the solvent from water to alcohol.

This survey will probably suffice for illustrating the convenience of the acidity and basicity constants as a measure of acid and basic strength. The change of the constants for the same acid and base passing from one medium to another is consistent with the terminology adopted, since an acid is said to possess the greatest strength in the medium, in which under equal conditions, i.e., at constant $\frac{c_A}{c_B}$, the highest acidity is produced. The definition

gives to no less degree a rational comparison between the strengths of acids and bases in unchanged medium, since the strength in this case may also be measured by the acidity or basicity in the solution under equal conditions.

Before the investigation is extended to comprise the actual dissociation phenomena of dissolved acids and bases, the important question of the acidity and the basicity of the medium itself has to be studied. A certain acidity and basicity will exist in most solvents without any addition of acid or base, and this property of the pure solvent is highly important in acid-base equilibria. If, for example, the acidity of water is defined similarly to the acidity of an acid dissolved in water, we shall have:

$$K_{\text{Acid. (H}_2\text{O)}} = \frac{c_{\text{OH}^-}}{c_{\text{H}_2\text{O}}} \cdot a_{\text{H}^+},$$

and this quantity might well be termed the *rational acidity constant* of water. It is seen, however, that the same exact meaning cannot be assigned to this constant, since $c_{\text{H}_2\text{O}}$ does not possess an explicit value as does the concentration of a dissolved substance. By analogy with the recognized custom of omitting the concentration of the molecules in the dissociation constant of a pure liquid, a new constant may be formulated as follows:

$$K_{\text{Acid (H}_2\text{O)}} = c_{\text{OH}^-} \cdot a_{\text{H}^+}$$

which has a definite value and which therefore in contradistinction to the preceding constant may be called the *conventional acidity constant* of water.

The same complication appears also in the definition of the acidity constant for the H_3O^+ -ion, since the corresponding base in this case is the pure solvent, the conditions being analogous also in the definition of the basicity constants. No difficulties are encountered, however, in the formulation of the corresponding formulas when the distinction between rational and conventional constants is maintained throughout the treatment.

The scheme, expressing the acid or basic function, together with the formulas for the two kinds of constants is given below. (Rational; conventional; acidity; and basicity are abbreviated to rat.; conv.; acid.; bas. respectively.)

The acid function for H_2O : $\text{H}_2\text{O} \rightarrow \text{OH}^- + \text{H}^+$

$$K_{\text{rat. Acid. (H}_2\text{O)}} = \frac{c_{\text{OH}^-}}{c_{\text{H}_2\text{O}}} \cdot a_{\text{H}^+} \quad K_{\text{conv. Acid. (H}_2\text{O)}} = c_{\text{OH}^-} \cdot a_{\text{H}^+}.$$

The base function for H_2O : $\text{H}_2\text{O} + \text{H}^+ \rightarrow \text{H}_3\text{O}^+$

$$K_{\text{rat. Bas. (H}_2\text{O)}} = \frac{c_{\text{H}_3\text{O}^+}}{c_{\text{H}_2\text{O}}} \cdot \frac{1}{a_{\text{H}^+}} \quad K_{\text{conv. Bas. (H}_2\text{O)}} = c_{\text{H}_3\text{O}^+} \cdot \frac{1}{a_{\text{H}^+}}.$$

The acid function for H_3O^+ : $\text{H}_3\text{O}^+ \rightarrow \text{H}_2\text{O} + \text{H}^+$

$$K_{\text{rat. Acid. (H}_3\text{O}^+)} = \frac{c_{\text{H}_2\text{O}}}{c_{\text{H}_3\text{O}^+}} \cdot a_{\text{H}^+} \quad K_{\text{conv. Acid. (H}_3\text{O}^+)} = \frac{a_{\text{H}^+}}{c_{\text{H}_3\text{O}^+}}.$$

The base function for OH^- : $\text{OH}^- + \text{H}^+ \rightarrow \text{H}_2\text{O}$

$$K_{\text{rat. Bas. (OH}^-)} = \frac{c_{\text{H}_2\text{O}}}{c_{\text{OH}^-}} \cdot \frac{1}{a_{\text{H}^+}} \quad K_{\text{conv. Bas. (OH}^-)} = \frac{1}{c_{\text{OH}^-} \cdot a_{\text{H}^+}}.$$

It follows from the definitions of the constants for conjugate acids and bases, that $K_{\text{Acid (H}_2\text{O)}} \cdot K_{\text{Bas. (OH}^-)} = K_{\text{Bas (H}_2\text{O)}} \cdot K_{\text{Acid (H}_3\text{O}^+)} = 1$, as may be seen from the above table both for the rational and conventional constants.

The corresponding expressions for other solvents may easily be

formulated. The conventional basicity constant in methyl alcohol is thus.

$$K_{\text{conv Bas (CH}_3\text{OH)}} = \frac{c_{\text{CH}_3\text{OH}_2^+}}{a_{\text{H}^+}}$$

In order to give these constants numerical significance a general standardization of the hydrion activity a_{H^+} has to be made. This might be achieved by assuming the absolute stoichiometric activity coefficient of the hydrion equal to unity in an infinitely dilute aqueous solution, which means that $c_{\text{H}_3\text{O}^+} = a_{\text{H}^+}$ in this solution. The conventional acidity constant of water is then identical with the "dissociation constant" in the classical sense, whereas the corresponding basicity constant for water is equal to unity.

About the constants, derived from this standardization for other solvents only little is known. The hydrion potential in ethyl alcohol is, according to the E.M.F. measurements of Larsson 0.149 volts higher than the potential in water of the same hydrion concentrations ($c_{\text{H}_3\text{O}^+} = c_{\text{C}_2\text{H}_5\text{OH}_2^+}$), i.e.:

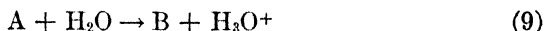
$$K_{\text{conv (Bas H}_2\text{O)}} = 330 K_{\text{conv Bas (C}_2\text{H}_5\text{OH)}}$$

Water is consequently from the conventional constants 330 times as strong a base as is ethyl alcohol, whereas the difference between the rational constants is somewhat less. Due to the existence of disturbing junction potentials, the figures thus obtained are of a considerable uncertainty.

It is of great importance to have these constants determined for different solvents. The general method for this purpose is to determine the hydrion activity in the two solvents after addition of known quantities of acid or base, sufficiently strong for producing in the solution the same concentration of the acid or base corresponding to the molecules of the medium. It might also be done by determining both the acidity or basicity constants and the dissociation constants of the acid and the base in the two solvents to be compared, as further explained below.

So far no attention has been paid to the actual "dissociation" of the acid and base, which is of such outstanding importance in

the classical theory. Proceeding now to an analysis of this process of dissociation it will be well to bear in mind that the free hydrion, as shown before, hardly has any chance of existing in the solution. The "dissociation" of an acid, when introduced in a pure solvent, is therefore dependent upon the basic character of the solvent, since the "dissociation" really means nothing else but transference of the hydrion from the acid to the solvent. The reaction may be written for the acid A and water as solvent as follows:

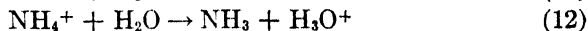
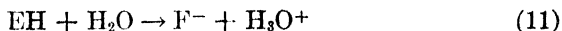


If the solvent completely lacks the character of a base, no "dissociation" will take place and the acid will dissolve unchanged in the medium.

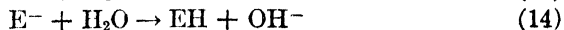
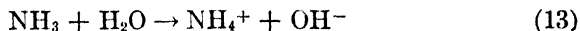
A reaction of the base similar to the above one (9) for the acid will take place only, when the solvent is able to yield the hydrion, necessary for the base if its basicity is to be manifested through a chemical reaction. In the case of water as a solvent the process is:



The schemes (9) and (10) give the characteristic acid and base equilibria, independent of the charges on A and B. In spite of the reactions for acetic acid and the ammonium ion in water being quite analogous:



the first is normally called a dissociation and the second an hydrolysis. Similarly of the two reactions:



it is usual to call (13) a dissociation and (14) a hydrolytic reaction. One and the same terminology ought, however, to be used for reactions like (11) and (12) which conform to scheme (9)—and again for (13) and (14) conforming to (10). It does not follow directly from this view, however, that schemes (9) and (10) ought

to be included in the same category, as has been the usual practice for acids and bases of the same charge.

A consideration of reaction (9) and analogous processes with the solvent shows that they do not represent a dissociation in the same sense of the word as applied to gaseous systems, but are in reality double decompositions, i.e., hydrolytic or solvolytic reactions. Nevertheless it seems adequate in the terminology of these reactions to disregard the solvent, the effect of which is the same in all cases, and call the process a dissociation in accord with the language historically recognized for neutral acids, the more so as the reaction in dilute solution is governed by laws conforming to the laws of a true dissociation. Once decided upon, this terminology has, however, to be carried through consistently in all cases, irrespective of the charge type. The formation of the H_3O^+ -ion should always be described as a dissociation even when the dissociating acid is for example the ammonium ion or another molecule carrying an electric charge.

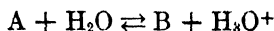
The application of the term dissociation to reaction (10) may be considered still more illogical, since the B-constituent does not decompose but on the contrary takes up a hydron. Nevertheless the similarity of the two schemes (9) and (10) both representing an acid-base equilibrium makes it desirable to retain the term "dissociation" also for the case of a base dissolved in an acid solvent.

We find it convenient therefore not only to keep the term dissociation for the reactions characteristic of acids and bases in appropriate solvents but also to generalize the application to comprise all such reactions, regardless of whether the acid or base is charged or electrically neutral. The term hydrolysis, which has hitherto been applied to some of these acid-basic reactions, is in this way eliminated, and the ambiguity of having two expressions for the same phenomenon cleared away. On the other hand it would perhaps be preferable, to characterize the dissociation as being hydrolytic, alcoholytic, aminolytic etc. according to the solvent involved in the process.

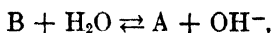
For characterizing various dissociation equilibria the usual dissociation constant employed for neutral acids and bases is well

suit. This constant will be denoted by K_A and K_B , for the acid and basic dissociations respectively. Here also, as for the acidity and basicity constant, the distinction between rational and conventional constants proves relevant, since the solvent participates in the process of dissociation.

If the acid and base dissociation be given by:



and



respectively, the mass action law yields on including the solvent:

$$\frac{c_B \cdot c_{H_3O^+}}{c_A \cdot c_{H_2O}} = K_{rat. A} \quad (15)$$

and:

$$\frac{c_A \cdot c_{OH^-}}{c_B \cdot c_{H_2O}} = K_{rat. B} \quad (16)$$

and by leaving out the solvent:

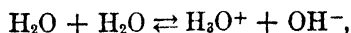
$$\frac{c_B \cdot c_{H_3O^+}}{c_A} = K_{conv. A} \quad (17)$$

and

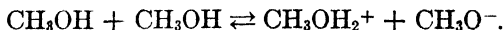
$$\frac{c_A \cdot c_{OH^-}}{c_B} = K_{conv. B}. \quad (18)$$

Without further specification K_A and K_B will be considered as denoting the conventional constants.

It follows from these considerations that an acid-base dissociation of the pure solvent can take place only if its character is amphoteric, i.e. if it possesses both acid and basic properties. The molecules of the solvent are then able both to give off and take up hydrions. For water the electrolytic dissociation is determined by the process:



one water molecule functioning as an acid forming OH^- and the other as a base forming H_3O^+ . The analogous process in methyl alcohol will be:



A low electrolytic dissociation for a pure solvent is hence due to the weak development of either its acid or its basic character. The great influence which the dielectric constant is known to exert upon the dissociation is included in its effect upon the acidity and basicity constants.

Contrary to the acidity and basicity constants for which no common measure exists the acid and basic dissociation constants will be equal for the pure solvent. The previous distinction between rational and conventional constants must also be maintained here so we can write:

$$\frac{c_{\text{H}_3\text{O}^+} \cdot c_{\text{OH}^-}}{c_{\text{H}_2\text{O}}^2} = K_{\text{rat A (H}_2\text{O)}} = K_{\text{rat B (H}_2\text{O)}}$$

$$c_{\text{H}_3\text{O}^+} \cdot c_{\text{OH}^-} = K_{\text{conv A (H}_2\text{O)}} = K_{\text{conv B (H}_2\text{O)}}$$

and analogously for methyl alcohol and other media of amphoteric character.

The question now arises, what is the significance of the dissociation constants as measures of strength in relation to the acidity and basicity constants which we have introduced above as a general measure in this respect.

If the equations for the acidity and basicity constants of the solvent are combined with the expressions for the dissolved acids and bases (1)–(2) and (15)–(18), we get:

$$K_A = K_{\text{Acid.}} K_{\text{Bas. (H}_2\text{O)}} \quad (19)$$

$$K_B = K_{\text{Bas.}} K_{\text{Acid (H}_2\text{O)}} \quad (20)$$

The dissociation constant of an acid is therefore the product of the acidity constant of the acid and the basicity constant of the medium, and the dissociation constant of a base is the product of the basicity constant of the base and the acidity constant of the

medium. Both K_A and K_{Acid} as well as K_B and $K_{Bas.}$ are equally well suited as measures of strength so long as the medium is kept constant. We note that K_A and K_{Acid} are altered to the same extent by the change from one definite medium to another, since the ratio K_A/K_{Acid} is according to (19) independent of the nature of the acid. The bases of course follow the same rule.

For the medium itself we may write:

$$c_{H_3O^+} c_{OH^-} = K_A (H_2O) = K_B (H_2O) = K_{Acid} (H_2O) \cdot K_{Bas.} (H_2O) \quad (21)$$

The acid and basic dissociation constants of the amphoteric solvent are consequently identical and equal to the product of its acidity constants as in the case of (19) and (20) for dissolved acids and bases. This equation is valid for the rational as well as for the conventional constant.

Most quantitative measurements of acid and basic strength have had as their aim the determination of K_A and K_B . In order to estimate the change in these constants from one solvent to another, the formulas (19) and (20) have to be consulted. In addition to the effect of the medium upon the acidity and basicity constants which has already been discussed, these equations show that the dissociation is determined by the acid and basic nature of the solvent. The changes exhibited by the acidity and basicity constants on alteration of the medium as compared with the corresponding changes in the dissociation constants are thus able to give information concerning the acid and basic character of the medium. Exact investigations for the purpose of determining the acid and basic nature of pure solvents are however not available, and due to this lack of experimental data the interpretation of the dissociation phenomena as regards their dependence upon the medium is rendered feasible only to a very limited extent.

Bjerrum, Unmack and Zechmeister (19) have carried out exact measurements on the dissociation of acetic acid and ammonia in methyl alcohol as well as on the dissociation constant of pure methyl alcohol. The acid dissociation constants, calculated from the constants given by Bjerrum, viz.: $K_{EH} = 10^{-9.65}$,

$K_{\text{NH}_4} = 10^{-5.92}$ and $K_{\text{CH}_3\text{OH}} = 10^{-17.01}$, are tabulated below together with the corresponding acid constants in water:

ACID	K_A IN METHYL ALCOHOL	K_A IN WATER	LOGARITHMIC CHANGE
Acetic acid	$10^{-9.65}$	$10^{-4.76}$	4.89
Ammonium ion	$10^{-11.09}$	$10^{-9.39}$	1.70

We see from these figures that the change from water to methyl alcohol causes the constants of both acids to decrease but much more for acetic acid than for the ammonium ion. According to the previous considerations a decrease in K_{Acid} for acetic acid and an increase for the ammonium ion have to be expected. The diminution in both constants is then explained on the basis of equation (19) by the smallness of the basicity constant of methyl alcohol in comparison with that of water. The great decrease in the dissociation constant observed in the case of acetic acid is explained by the effects of the change of medium upon K_{Acid} and $K_{\text{Bas (Medium)}}$ operating in the same direction, whereas the far smaller decrease in the case of the ammonium ion is due to the two effects counteracting each other.

The change in acidity constant from water to 95 per cent methyl alcohol has been determined by Mizutani (142) for a series of acids, his values being easily extrapolated to 100 per cent methyl alcohol. Acetic acid shows a decrease of 2.3 and ammonium ion an increase of 0.8 logarithmic units. From this change combined with the above change in the dissociation constants the following values are calculated by equation (19):

Change in constants from water to methyl alcohol

	$\Delta \log K_A$	$\Delta \log K_{\text{Acid}}$	$\Delta \log K_{\text{Bas (Medium)}}$
Acetic acid	-4.9	-2.3	-2.6
Ammonium ion	-1.7	+0.8	-2.5
Average			-2.55

The values obtained from the two independent computations are thus in good agreement, indicating that pure methyl alcohol is about 350 times as weak a base as is pure water.

An analogous calculation, although with greater uncertainty, can be carried out for ethyl alcohol. The basicity constant relative to that of water is found by comparison of the dissociation constants and acidity constants of acids in water and alcohol as exhibited in the following table. Δ is the increase from water to alcohol. The figures of the first column are taken from a survey by Larsson (125) and those in the second column from Michaelis and Mizutani (138).

Change in constants from water to ethyl alcohol

	$\Delta \log K_A$	$\Delta \log K_{Acid}$	$\Delta \log K_{Bas} \text{ (Medium)}$
Acetic acid	-5.5	-2.7	-2.8
Benzoic acid	-5.9	-3.2	-2.7
Salicylic acid	-5.7	-2.8	-2.9
Average			-2.8

According to these figures ethyl alcohol is about 600 times as weak a base as water. The value 330 found previously from Larsson's measurements are in moderate agreement with this result. If Larsson's figures are right methyl and ethyl alcohol should have practically the same basic strength.

It is now possible to calculate the acidity constants of the alcoholic media by means of formula (21). Thus for methyl alcohol:

$$\frac{K_{A(H_2O)}}{K_{A(CH_3OH)}} = \frac{K_{B(H_2O)}}{K_{B(CH_3OH)}} = \frac{K_{Acid(H_2O)} \cdot K_{Bas(H_2O)}}{K_{Acid(CH_3OH)} \cdot K_{Bas(CH_3OH)}}$$

or by introducing numerical values:

$$\frac{K_{Acid(H_2O)}}{K_{Acid(CH_3OH)}} = \frac{10^{-14.1}}{10^{-17.0}} \cdot 10^{-2.5} = 10^{0.4} = 2.5$$

By using Danner's (41) value for the dissociation constant of ethyl alcohol at 18°C.:

$$K_{A(C_2H_5OH)} = 10^{-19.1}$$

we get:

$$\frac{K_{Acid(H_2O)}}{K_{Acid(C_2H_5OH)}} = \frac{10^{-14.1}}{10^{-19.1}} \cdot 10^{-2.8} = 10^{2.2} = 160.$$

If the data—used for these calculations—be trustworthy, the acidity would change only slightly from water to methyl alcohol, but considerably from methyl alcohol to ethyl alcohol. This result seems somewhat improbable. The experimental results are, therefore, at least to some extent not satisfactory.

Further applications of the ideas developed in this chapter can be made on the basis of the important and extensive material produced by H. Goldschmidt and his pupils. The detailed treatment, necessary for investigating these results, are somewhat beyond the scope of this treatise, however, and cannot therefore be undertaken here.

VIII. AN EXTENDED THEORY OF ACID AND BASIC CATALYSIS

The change in the conception of acids and bases as developed in the two preceding chapters is likely to imply also a change in our views of the kinetic phenomena exhibited by this group of substances. In fact the subjection of the general thermodynamic properties of acids and bases to such an elaborate treatment was made partly because the kinetic significance of these general properties was to be expected. It might even be said that it is just in the field of catalysis that the effect of the reformulation of the acid-basic function manifests itself in a particularly definite and obvious manner.

That the hydrion in solution exists only in the solvated state is a fact of paramount importance in elucidation of the intrinsic nature of acid catalysis. It follows therefrom that hydrion catalysis in aqueous solution cannot be due to the free hydrion but has rather to be ascribed to the H_3O^+ or oxonium ion. However, this ion being an acid, which according to our previous statements cannot claim a unique position within the group of acids, there seems to be no a priori reason to try to reserve for it a unique position as an acid catalyst. On the contrary it would be logical to attribute the property of acid catalysis to all acid molecules, i.e. to all molecules capable of giving off a hydrogen nucleus. The effect of the various acids should then depend upon the readiness with which the nucleus is split off, this property being in its turn dependent upon the strength of the acid.

This mode of reasoning is valid, if the catalyzed reaction is not subject to a specific effect. However, the possibility of a reaction being catalyzed by the acid XH and not by YH must not be overlooked. The transformation of acetochloroamino-benzene into p-chloroacetanilide is thus catalyzed by hydrochloric acid, but not by nitric acid. Such specific phenomena are certainly of considerable interest but since we are going to develop the general theory of acid-basic catalysis, it will not be necessary to deal with them here.

The conclusions for basic catalysis will be quite analogous. Since the hydroxyl ion does not, in principle, assume any unique position among the bases, basic catalysis cannot be attributed solely to this particular ion. In general basic catalysis must be looked upon as the effect of the various bases involved and is therefore to be correlated with the ease with which a hydrogen nucleus is taken up, i.e. to the strength of the base.

The extended theory here outlined thus removes the hydrion and hydroxyl ion from the unique position held by them in the classical theory of dissociation and catalysis. It ascribes quite generally acid and basic catalysis directly to molecules of acid and basic character. On the basis of these ideas a number of questions of importance in catalysis arise, some of which will be discussed in the present chapter. First we shall have to consider however the new theory in its relation to the dual theory (Chap. 2), with which it might seem to have some points in common through the importance attached to the "undissociated acids" by both theories.

There is no doubt as to the similarities between the contents of the two theories, since the dual theory as well as the extended catalytic theory supposes the "undissociated acids" to exert catalytic effect. The former, however, considers the "hydrion" and the undissociated acid molecule as quite different species, thus setting up consciously a distinction between acid catalysis and hydrion catalysis. The present theory avoids such a dualism, looking upon hydrion catalysis simply as a particular instance of general acid catalysis. This consistency in the way of regarding catalytic phenomena is the theoretical strength of the extended theory of catalysis. The dual theory on the other hand does not

rest upon theoretical considerations and is—as pointed out in Chapter 2—based more than anything else upon misinterpretation of the experimental facts. The catalytic effect used as evidence for the dual theory is in most of the cases caused by phenomena which are foreign to the views inherent in the dual theory. The foundation of the two theories are therefore entirely different, even if their contents exhibit a certain similarity.

Except for those presented in the dual theory no views have been advanced previous to the present theory of catalysis, which might render possible a similar explanation of catalytic phenomena. Still mention must be made, of an investigation by Fajans, (61) suggested to him by Bredig, in which the decomposition of camphor and bromocamphor-carbonic acid in presence of nicotine, quinine, quinidine, etc. is interpreted as a certain kind of basic catalysis. The reaction was however, explained, as being a decomposition of the "salt" formed with the base, i.e., more or less in conformity with the ionization theory of catalysis. The author mentions that a deeper insight into the reaction might be obtained when the investigation is extended to comprise a greater number of bases and consideration of their strength. An extension of our general ideas in catalysis is however not afforded by this work, its aim being rather to examine the stereochemical characteristics of the catalyst.

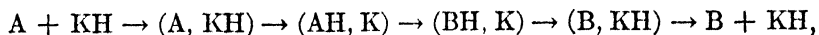
In some relation to this work Ljunggren (130) investigated the catalytic decomposition of acetoacetic acid in presence of several organic bases. His results are of considerable interest, but no new points of view for their interpretation appear to have been obtained.

Investigations of this nature may, however, prove of great value for the development of the present theory of catalysis. For example, this is the case in the mutarotation reaction of nitrocamphor in various solvents, as studied by Lowry (135), in which substances like piperidine, aniline, sodium ethylate, water, etc. showed great catalytic effect. It is probable that catalysis by basic molecules takes place, but the experimental data do not suffice for stating exactly the character of the reaction.

In later investigations on mutarotation by Lowry and his co-workers (65, 132, 134, 136), attention has especially been called

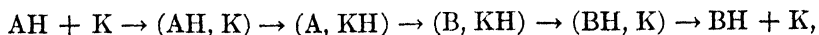
to the acid or basic character of the solvent. Evidence of much greater consequence for the extended theory of catalysis has been furnished, however, by certain reactions in aqueous solutions especially by the catalytic decomposition of nitramide and the mutarotation of glucose. Before proceeding to present the experimental verifications of the theory obtained by means of these reactions we shall try to analyze the mechanism of catalysis in order to obtain some idea of the important question of the significance of the acid and basic strength of the catalyst.

The primary condition for catalysis is the possibility of the molecules of the catalyst and the substrate forming new molecular complexes on colliding. These complexes may in conformity with previous proposals be called critical complexes. If the catalyst be an acid it seems natural to picture the reaction in the first step as a transference of the hydrogen nucleus of the catalyst into the substrate molecule. The binding of the hydrogen nucleus occurring in molecules of high energy content, might then cause a disruption of the structure of the substrate molecule resulting in the formation of a stable product as the next step. The catalytic mechanism of the reaction $A \rightarrow B$ might then be presented as follows:



where KH is the acid catalyst and the parentheses signify the molecular complexes reacting.

Similarly in basic catalysis the shock, initiating the reaction, is supposed to be caused by the transference of a hydrogen nucleus from the substrate to the catalyst within the complex. As an illustration of the mechanism of the process $AH \rightarrow BH$ we thus can write:



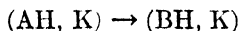
where K is the catalyst base.

This picture can, of course, only give a rough sketch of what happens in the process, since the symbols in these schemes are rather undefined in their character. For example, nothing is

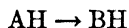
said as to whether the passage of the hydrogen nucleus proceeds continuously or whether, as may be expected, it occurs as a quantum jump. Also the effect of the solvent molecules in acting as receivers and transmitters of energy has been left out of the considerations. Nevertheless the above schemes, however incomplete in presenting an exact picture of catalysis, seem to be serviceable as a basis for further conclusions in particular as regards the influence of the acid and basic strength.

The mechanism suggested contains an important consequence with respect to the nature of the substrate. It is obvious that acid and basic catalysis require a substrate of basic and acidic nature, respectively. These properties need not be developed, however, to such an extent as to permit of detection by general methods for measuring acidity and basicity. Only in the case of substrates and solvents containing no hydrogen can the absence of acid character and therefore of basic catalysis with certainty be predicted.

From the above one might perhaps feel inclined to believe that after all the distinction between the ionization theory and the present theory of acid and basic catalysis is comparatively slight. For instance the kinetic reaction determining the velocity in the acid catalytic transformation $A \rightarrow B$, according to the extended theory is:

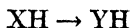


and according to the ionization theory:

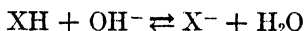


In the latter theory, however, AH is a normal molecule whose conversion into BH is no more intelligible than the original reaction $A \rightarrow B$ for which it was expected to furnish an explanation, whereas (AH, K) is a critical complex capable of instantaneous reaction. Furthermore, as pointed out in Chapter 3 the empirical consequences of the two points of view may differ fundamentally, which makes it possible to take a more definite standpoint towards the main idea of the ionization theory.

For this purpose consider the following process catalyzed by bases:

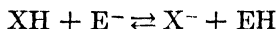


As mentioned in Chapter 3, it is impossible from the shape of the reaction curve to make any distinction between a hydroxyl ion catalysis and a spontaneous reaction of the X⁻-ion. This is due to the proportionality between the concentration of the X⁻-ion and the product of the XH and OH⁻-concentrations in the equilibrium:



This proportionality in turn is conditioned by the fact that the fourth molecule in the equilibrium, the acid conjugate to the catalyzing base, is present, as the solvent, i.e. in unvariable concentration.

If the hydroxyl ion is present in non-aqueous solution, or if the basic catalysis in water is due to some catalyst other than the hydroxyl ion this difficulty obviously disappears. If the acetate ion, for example, be the catalyst the corresponding equilibrium:



evidently permits of a change in the product of the XH and E⁻-concentrations without requiring a simultaneous change in the concentration of the X⁻-ion. It should be possible from that kind of experiments to decide, which of the two theories is correct. For instance in acetate--acetic acid buffers of constant ratio between the buffer constituents but of variable total concentration, the ionization theory in the absence of salt effects demands constant rate of reaction whereas the extended theory of catalysis would require variations if a detectable catalytic effect can be ascribed to the acetate ion.

It should be noted, however, that even if the experimental test comes out in favor of the extended theory the possibility of a collateral spontaneous decomposition of the X⁻-ion in the course of the OH⁻-catalysis is not excluded. In order to study this possibility a closer investigation of the general laws governing acid

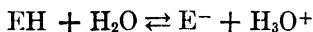
and basic catalysis particularly as regards the effect of the strength of the catalyst is required.

The dependence of the catalytic effect upon the strength of the catalyzing acids and bases is actually a question of paramount importance in the present theory. The readiness with which a hydrogen nucleus is given off by an acid and taken up by a base may be assumed to be greater the greater the strength of the acid and the base, and hence the catalytic effect would be expected to increase with increasing strength of the acid and the base. The simplest scheme would be to assume direct proportionality between the catalytic constant and the strength constant of the acid or basic catalyst, in which case:

$$k_a = GK_A,$$

where k_a and K_A denote the catalytic and dissociation constant of the acid catalyst respectively and G is a constant dependent only upon temperature, pressure, medium and substrate. Obedience to such a simple relation, however, cannot be expected for the following reasons:

The probability of the hydrogen nucleus leaving the acid molecule and forming the reacting complex with the substrate can hardly be directly related to the strength of the acid but rather to the *rate* at which the acid dissociates, i.e. reacts with the solvent to form hydrion solvate and the conjugate base. The "dissociation" equilibrium in an acid solution is, like any other equilibrium, supposed to be kinetic in nature; so that the equilibrium for, say, acetic acid in water:



is the state in which the two opposing reactions proceed with the same velocity. If the reaction from left to right be (arbitrarily) called a dissociation, from right to left an association, and the corresponding velocity constants $k_{\text{diss.}}$ and $k_{\text{ass.}}$ the dissociation constant is given by:

$$K_A = \frac{k_{\text{diss.}}}{k_{\text{ass.}}} \quad (1)$$

It is obvious, then, that the same strength may correspond for different acids to quite different values of $k_{\text{diss.}}$ and $k_{\text{ass.}}$.

We must expect that the catalytic constant k_a is more closely related to $k_{\text{diss.}}$ than to K_A because the processes of catalysis and of dissociation are of analogous character. In the former reaction the hydrion is given off to a complex, which undergoes a further decomposition, whereas in the latter the hydrion is given off to one of the solvent molecules, with which it might stay united for a certain length of time before being lost again. However, what happens after addition of the hydrion cannot be essential for the probability of the addition process and it is therefore most likely that the velocity of the two processes follows the same law.

This conclusion may be further defined by introducing the critical energy of reaction. The difference between the critical energies $E_{\text{diss.}}$ and E_a corresponding to the "dissociation" and the catalysis, is independent of the catalyst for a given reaction in the same solvent. If the velocity constant for the acid dissociation be $k_{\text{diss.}}$ and for the catalysis k_a we get:

$$k_a = k e^{-\frac{E_a}{RT}}$$

and

$$k_{\text{diss.}} = k e^{-\frac{E_{\text{diss.}}}{RT}}$$

whence the ratio between the two velocity constants is found equal to:

$$\frac{k_a}{k_{\text{diss.}}} = e^{-\frac{E_a - E_{\text{diss.}}}{RT}}$$

This equation shows that for different catalysts in the same reaction and medium the velocity constants of the acid catalysis and the acid dissociation are proportional.

Experiments for determining the velocity of dissociation have shown it to be of a high order of magnitude without having succeeded in determining absolute values. As to the relative dissociation velocities we may, however, make certain conjectures.

According to the kinetic conception of the dissociation constant as expressed by formula (1), this constant can only vary simultaneously with $k_{\text{diss.}}$ and $k_{\text{ass.}}$. The simplest assumption is that—in any case for acids of the same type—an increase in the rate of dissociation from one acid to another is accompanied by a decrease in the rate of association and vice versa, so that in other words the change in strength, observed for a series of analogous acids, is caused by concerted changes in $k_{\text{diss.}}$ and $k_{\text{ass.}}$.

This means that passing on from one acid to another K_A will have to change more markedly than do $k_{\text{diss.}}$ and $k_{\text{ass.}}$. This result may be expressed by the following equation:

$$k_{\text{diss}} = G K_A^x \quad (2)$$

in which G is a constant and x a proper fraction. The numerical value of the latter cannot be estimated a priori, nor is it permissible to assign in advance any constant value to it, independent of the acid strength. It is however reasonable to assume x to be a continuous function of K for acids of the same type.

If the proportionality between the velocities of dissociation and of catalysis is introduced into equation (2) we get:

$$k_a = G_1 K_A^x \quad (3)$$

as the relation which may be anticipated to exist between the catalytic constant and the dissociation constant of an acid.

Analogous considerations can be applied to basic catalysis. We assume proportionality between the catalytic constant and the velocity constant of association and further that the ratio between these constants does not change with the nature of the base. Furthermore from formula (1) and (2) is deduced:

$$k_{\text{ass}} = G K_A^{x-1}$$

where K_A is the dissociation constant of the corresponding acid. If we now introduce the base constant from the relation:

$$K_A K_B = \text{dissociation constant of medium}$$

we get:

$$k_{\text{ass.}} = G' K_B^{1-x} \quad (4)$$

and consequently:

$$k_b = G_2 K_B^{1-x}. \quad (5)$$

It follows from these equations first, that the value of x is independent of the nature of the substrate. Secondly the equations give information concerning the relation between the catalysis of conjugate acids and bases. It follows from the derivation that x has the same value in the two equations for conjugate acids and bases:

$$k_a = G_1 K_A^x, \quad k_b = G_2 K_B^{1-x}$$

If x —as a proper fraction—varies from 0 to 1 the sensitivity of the acid catalysis to changes in K_A will rise gradually from 0 to 1, while at the same time the sensitivity of the basic catalysis will fall gradually from 1 to 0. A sensitivity equal to zero means that the catalytic constant does not change with the strength of the catalyst, whereas a sensitivity equal to unity means that the variation is directly proportional to it. These two cases are of course limiting cases, the sensitivity normally deviating from these minimum and maximum values. However, it appears from the formula, that if the acid catalysis displays great sensitivity to changes in the strength constant, the basic catalysis will display low sensibility and vice versa.

It is obvious that the views advanced here lead to a theory of acid and basic catalysis, which in its principle rests on a broader foundation than previous theories. It must be admitted, however, that the premises which lie at the root of the theory, although plausible, are not absolutely cogent in their nature, and it might therefore prove necessary on closer examination to modify these assumptions in different directions. Experimental test is of course the most direct way of attaining a measure of the value of the theory.

In the experimental investigation of the present theory of catalysis we have, however, to consider some particular circumstances of significance in correlating the efficacy of catalysts of different type with the corresponding strength constants. The catalytic as well as the acid and basic properties are generally not

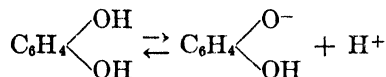
determined solely by the firmness with which the hydron is attached to the acid molecule, but also depend upon the number of ionizable hydrogen nuclei possessed by the acid molecule and upon the number of points in the base molecule to which the hydron can attach itself.

If the acid molecule contains only one acid hydrogen atom and the corresponding base only one point of attack, we have the statistically simplest acid-base system, represented for instance by the phenol-phenolate system:

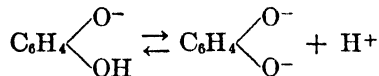


The number of acid hydrogen atoms in the acid molecule being denoted by p and the number of points of attack in the base molecule by q , we have for this simple case: $p = q = 1$.

In the first dissociation step for hydroquinone:



we have two ionizable hydrogen atoms in the acid and one point of attack in the base. Hence $p = 2$ and $q = 1$ in this case, whereas for the second dissociation step:



$p = 1$ and $q = 2$, etc.

For the purpose of studying the influence of these statistical conditions upon acid and basic strength we must attribute to all ionizable hydriions the same fixity in the molecule. Of course we must also assume the statistical conditions to be completely obeyed according to the above simple picture. It is then clear that in the above examples hydroquinone, in the first step of dissociation will be twice, and in the second step half as strong an acid as phenol, and that, in general, the dissociation constant of an acid will be given by:

$$K_A = \frac{p}{q} K_{A_0} \quad (6)$$

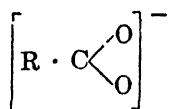
where K_{A_0} signifies the dissociation constant in the statistically simple acid-base system of the same hydrogen fixity. The ratio p/q will be called the statistical acid strength factor. For the conjugate base the relation:

$$K_B = \frac{q}{p} K_{B_0} \quad (7)$$

analogous to (6) is easily seen to hold, the ratio q/p being termed the statistical basic strength factor. The following table gives the statistical factors for a series of acid-base systems as computed from the formulas (6) and (7). The systems are arranged in the order of decreasing values of p/q .

ACID	BASE	p	q	p/q
$\text{Cr}(\text{H}_2\text{O})_6^{+++}$	$\text{Cr}(\text{H}_2\text{O})_5\text{OH}^{++}$	12	1	12
NH_4^+	NH_3	4	1	4
$\text{C}_6\text{H}_5\text{NH}_3^+$	$\text{C}_6\text{H}_5\text{NH}_2$	3	1	3
OH_3^+	H_2O	3	1	3
$\text{Co}(\text{NH}_3)_5\text{H}_2\text{O}^{+++}$	$\text{Co}(\text{NH}_3)_5\text{OH}^{++}$	2	1	2
$\text{Cr}(\text{NH}_3)_2(\text{H}_2\text{O})_2(\text{OH})_2^+$	$\text{Cr}(\text{NH}_3)_2(\text{H}_2\text{O})(\text{OH})_3$	4	3	4/3
H_3PO_4	H_2PO_4^-	3	2	3/2
$\text{C}_2\text{O}_4\text{H}_2$	$\text{C}_2\text{O}_4\text{H}^-$	2	2	1
$\text{C}_6\text{H}_5\text{NH}^+$	$\text{C}_6\text{H}_5\text{N}$	1	1	1
$\text{C}_6\text{H}_5\text{N}(\text{CH}_3)_2\text{H}^+$	$\text{C}_6\text{H}_5\text{N}(\text{CH}_3)_2$	1	1	1
H_2PO_4^-	HPO_4^{--}	2	3	2/3
CH_3COOH	CH_3COO^-	1	2	1/2
$\text{C}_2\text{O}_4\text{H}^-$	$\text{C}_2\text{O}_4^{--}$	1	4	1/4
HPO_4^{--}	PO_4^{---}	1	4	1/4

It should not be overlooked, however, that these values are not entirely free of assumptions. For example, it has been assumed that the carboxyl group after the ionization of its hydrogen atom has two points of attack; in other words the two oxygen atoms of the ion:



are equally available for the hydrion. This implies that the nega-

tive charge is not specifically attached to the oxygen atom in the hydroxyl group and that the formula $\text{R}-\text{C} \begin{smallmatrix} \text{O} \\ \parallel \\ \text{OH} \end{smallmatrix}$ is justified only

as a description of the undissociated acid. Looking upon this condition from the standpoint of structural chemistry such cases may however be of special interest since the conclusions regarding statistical factors drawn from catalytic experiments might in many cases throw some light on difficultly accessible questions of structural chemistry.

For acid and basic catalysis the significance of the statistical strength factors is seen in the following manner. The previous deduction of the relation between the catalytic and strength constants:

$$k_a = G_1 K_A^x, \quad k_b = G_2 K_B^{1-x}$$

pre-supposes the simple statistical condition $p = q = 1$. These expressions ought now to be written:

$$k_{a_0} = G_1 K_{A_0}^x, \quad k_{b_0} = G_2 k_{B_0}^{1-x} \quad (8), (9)$$

in order to distinguish between the simple and the more complicated cases. These expressions are also valid for the more complicated cases as far as the influence of firmness of attachment of the hydrion upon the catalysis is concerned. However, an additional factor enters, namely that the catalytic effect of an acid may be proportional to the number of acid hydrions and the effect of a base proportional to the points exposed to attack from the hydrion. Therefore:

$$k_a = p k_{a_0}, \quad k_b = q k_{b_0}$$

Introducing these equations together with (6) and (7) into (8) and (9) the following expressions for the acid catalysis are obtained:

$$k_a = p G_1 K_{A_0}^x = p G_1 \left(\frac{q}{p} K_A \right)^x$$

or

$$k_a = G_1 K_A^x q^x p^{1-x} \quad (10)$$

and for the basic catalysis:

$$k_b = qG_2K_{b_0}^{1-x} = qG_2 \left(\frac{p}{q} K_B \right)^{1-x}$$

or

$$k_b = G_2 K_B^{1-x} q^x p^{1-x} \quad (11)$$

These equations require, just as in (6) and (7) a complete conformity to the statistical requirements.

As the chief result of these considerations we may state that the presence of complicated statistical conditions in the acid-base system will affect the catalytic activity of the acid and the base in proportion to the change in the factor:

$$q^x p^{1-x}$$

this effect according to formula (10) and (11) being the same in acid and basic catalysis. If it were permissible to consider x as independent of the nature of the reaction it would follow from the results of the nitramide catalysis that x is considerably smaller than $1-x$ and that the factor p therefore is of higher influence than is q . In this case the catalysis by bases would depend upon their strength constants to a greater extent than would catalysis by acids.

In the special case of hydrion catalysis in aqueous solutions p equals 3, since H_3O^+ acts as a catalyst, whereas p , for example, for acetic acid and related organic acids is equal to unity. This circumstance of course to some extent displaces the catalytic conditions in favor of the hydrion more or less obscuring the catalytic effect of the weak acid.

Investigations on the catalytic decomposition of nitramide have furnished the most important experimental evidence of the present theory of acid and basic catalysis. This substance discovered by Thiele and Lachmann (174) corresponds to the formula $H_2N_2O_2$ without showing the chemical properties expected of the amide of nitric acid. As already found by the discoverers, the

substance is decomposed practically instantaneously in basic solutions according to the scheme:



A closer investigation (37, 34) of this reaction in aqueous solutions has shown that within certain limits of basicity the process is composed of two partial reactions; a spontaneous one, proceeding independently of the acidity of the solution, and a catalytic one. *All bases seem to act as catalysts*, the term base being used in conformity with the definition in Chapter 6. If the solution be sufficiently acid, the base effect is depressed, and the spontaneous reaction is the only factor of importance. This reaction is strictly monomolecular and has a half-period of 13.7 hours at 15°.

In an 0.01 equimolar acetate buffer solution (0.01 m. acetate + 0.01 m. acetic acid) the rate of decomposition is about 15 times as great as the velocity of the spontaneous reaction. This effect as may easily be shown by varying the ratio acetate/acetic acid, is proportional to the concentration of the acetate ion and should therefore be considered as an acetate ion catalysis. The anions of other organic acids such as propionic acid, benzoic acid, formic acid, salicylic acid, etc. have a similar effect. The same behavior is furthermore shown by anions of polybasic acids and by amines like aniline, toluidine, chloroaniline, etc. Since changes in hydrion concentration at constant concentration of anions or of amines do not affect the catalysis, the reaction conforms actually to the requirements of the extended theory of basic catalysis.

The determination of the constants in the catalytic decomposition of nitramide when brought into relation with the corresponding strength constants yields the following expression:

$$k_b = 6.2 \cdot 10^{-5} K_B^{0.83},$$

which is valid for basic anions of a series of organic acids, whereas the expression:

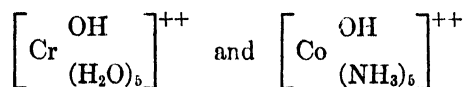
$$k_b = 1.7 \cdot 10^{-4} K_B^{0.75}$$

holds for cyclic amines.

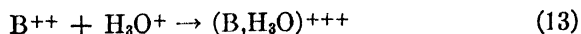
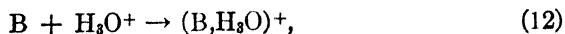
These equations conform to formula (5), deduced from theoretic-

cal considerations, the exponent $1 - x$ being a proper fraction while the above theoretical reasoning was unable to yield any information as to the variation of the exponent with the strength of the base, the measurements give the interesting result that $1 - x$ is constant within every group of bases, despite the considerable changes (from $2 \cdot 10^{-6}$ to 10^{-2} for the anions and from $0.7 \cdot 10^{-6}$ to $2 \cdot 10^{-3}$ for the amines) in the dissociation constants.

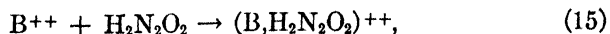
While the effect of the cyclic amines and the anions of the weak acids are but slightly different other basic catalysts exhibit a much higher catalytic activity. The polyvalent basic cations, as for example:



possess extraordinary catalytic effects easily noticeable even in strongly acid solution. The difference in the catalytic efficiency of different catalytic groups is not surprising, but on the contrary must be expected from a consideration of the catalytic mechanism. Since the association is a reaction between the base molecule and a positively charged hydrion, whereas the catalytic reaction takes place between the base molecule and a neutral nitramide molecule, parallelism between k_b and K_b is only to be expected as long as the catalyst is of unchanged charge type. If we consider the association and catalytic processes for two bases of different charge: B and B^{++} :



and



and assume the velocity to be the same for (14) and (15), i.e. B and B^{++} having the same catalytic influence, reaction (12) must necessarily proceed faster than (13) owing to the marked repulsive

forces affecting the latter reaction. In the case of the same catalytic effect B is consequently a stronger base than B^{++} , i.e. at the same rate of association—and this in turn, according to our theory, means the same basic strength—the catalytic effect is highest for the ion carrying the highest positive charge. This explains the high catalytic power of the bivalent hydroxo ions, but a similar effect called for by the theory when comparing bases having one negative and zero charge respectively does not seem to be present.

The problems of great physico-chemical interest further attached to the nitramide catalysis have been treated in the papers quoted above and need not be taken up here. The factors there discussed should, however, be taken into consideration in a general quantitative examination of the theory.

Another reaction probably of no less significance for importance in the theory of acid and basic catalysis is the mutarotation of glucose. Euler has found that certain ions have an anomalous effect upon the mutarotation of glucose, and Kuhn and Jacob have subjected this anomaly to a closer study in their paper previously mentioned. The phenomenon observed is the acceleration of the glucose transformation by certain ions like acetate, phosphate and citrate ions. The latter authors interpret the phenomenon as being an effect of these ions upon the activity of the reactants, this interpretation being obviously related to the view of the general significance of activity in catalytic reactions which has already been criticized in chapter 4. This explanation of the specific effect of the ions however offers no clue to the understanding of the phenomenon.

It is natural, however, to apply to this "anomaly" the same ideas which have proved so fruitful in the case of the nitramide catalysis. The reaction is a transformation of alpha-glucose into beta-glucose leading to an equilibrium between the two modifications (155). Hydrions and hydroxyl ions as well as acetate, phosphate and citrate ions act as catalysts. Furthermore the reaction is partly a "spontaneous" one; i.e. within a certain range of acidity the velocity is practically constant in absence of other catalysts.

All of these phenomena are undoubtedly to be interpreted on the basis of the extended theory of acid and basic catalysis.

Examination of this hypothesis by means of the data of Kuhn and Jacob is made difficult by the fact that these authors have conducted their experiments from the standpoint of the "activity theory" without, however, carrying through the work consistently on this basis. It is possible, however, by introducing concentration instead of activity throughout to obtain some idea of the magnitude of the catalytic effect.

The following equation for the spontaneous reaction by hydrions and hydroxyl ions according to Kuhn and Jacob is valid at 25°C.:

$$10^4 k = 104 + 3300 c_{H^+} + 9.3 \cdot 10^7 c_{OH^-}$$

This equation does not differ much from the formula given by Hudson mentioned previously. In the experiments with acetate, phosphate, etc. that part of the reaction, not due to the presence of these anions, can be eliminated by this equation, c_{H^+} and c_{OH^-} being approximately known, and in this way the effect of the anions can be calculated approximately. This calculation shows primarily that the effect is nearly proportional to the concentration of the anion in agreement with the theory. Moreover the catalytic effect is found to increase with the basic strength of the anion as in the case of the nitramide catalysis; but scarcity of data does not permit a determination of the quantitative relation between catalytic constant and basic strength.

However, in a study of mutarotation at present in progress in this laboratory, we have been able to establish definitely the presence of such catalysis and to determine it quantitatively for a series of bases of varying type. Similarly—and in this respect the results extend further than those for nitramide—we have obtained an analogous effect for acids in agreement with the views given here. Experiments on mutarotation have therefore already contributed importantly to the establishment of the extended theory. When we consider that these new phenomena have been observed in the study of a reaction which for a generation has been the subject of extensive experimental research, we may well believe

that a systematic study in the field of acid and basic catalysis will afford corresponding results in illustration of the ideas developed in these chapters.

LITERATURE

- (1) ÅKERLOF, *Z physik Chem.* **98**, 260 (1921).
- (2) ACREE, *Am Chem J* **39**, 513 (1908).
- (3) ACREE, *Am Chem. J* **48**, 352 (1912).
- (4) ACREE, *Am Chem J.* **49**, 345 (1913).
- (5) ACREE A. JOHNSON, *Am. Chem J* **37**, 410 (1907).
- (6) ACREE A. JOHNSON, *Am Chem. J.* **38**, 258 (1907).
- (7) ACREE A. NIRDLINGER, *Am Chem. J* **38**, 489 (1907).
- (8) ARRHENIUS, *Bih Stockh Akad* **8**, Nr. 14 (1884), *Z. physik. Chem.* **1**, 631 (1887).
- (9) ARRHENIUS, *Z physik. Chem.* **1**, 111 (1887).
- (10) ARRHENIUS, *Z physik Chem* **2**, 284 (1888).
- (11) ARRHENIUS, *Z physik Chem* **4**, 226 (1889).
- (12) ARRHENIUS, *Z physik. Chem.* **5**, 1 (1890)
- (13) ARRHENIUS, *Z. physik Chem.* **31**, 197 (1899).
- (14) AUERBACH, *Ergb. exakt Naturwiss* **1**, 228 (1922).
- (15) BAUR, *Ann* **296**, 95 (1897)
- (16) BAUR, *Z Elektrochem* **11**, 936 (1905).
- (17) BJERRUM, *Z Elektrochem.* **24**, 321 (1918), *Fys Tidsskr.* **15**, 59 (1916)
- (18) BJERRUM, *Z physik Chem.* **108**, 82 (1924)
- (19) BJERRUM, UNMACH O ZECHMEISTER, *D. Vid Selsk Medd.* **5**, Nr 11 (1924).
- (20) BLANKSMA, *Rec. P B* **21**, 366 (1902), **22**, 290 (1903)
- (21) BORN, *Z. physik.* **1**, 45 (1920)
- (22) BRAY A. LIVINGSTON, *J. Am. Chem Soc.* **45**, 1251 (1923)
- (23) BREDIG, *Z Elektrochem* **18**, 534 (1912).
- (24) BREDIG U. FRAENCKEL, *Ber* **39**, 1756 (1906)
- (25) BREDIG U RIPLEY, *Ber* **40**, 4015 (1907)
- (26) BREDIG U. WALTON, *Z physik Chem* **47**, 185 (1904).
- (27) BRONSTED, *J Chem Soc.* **119**, 574 (1921)
- (28) BRÖNSTED, *Z physik. Chem* **102**, 169 (1922)
- (29) BRONSTED, *J Am Chem. Soc* **44**, 938 (1922)
- (30) BRONSTED, *Rec P B* **42**, 718 (1923).
- (31) BRONSTED, *Z physik. Chem* **115**, 337 (1925).
- (32) BRÖNSTED, *J Phys. Chem* **30**, 777 (1926).
- (33) BRONSTED U DELBANCO, *Z anorg. allgem. Chem* **144**, 248 (1925).
- (34) BRONSTED U. DUUS, *Z physik Chem* **117**, 299 (1925).
- (35) BRONSTED A. KING, *J. Am. Chem. Soc.* **47**, 2523 (1925).
- (36) BRÖNSTED A. LAMER, *J Am Chem Soc* **46**, 555 (1924).
- (37) BRONSTED U. PEDERSEN, *Z. physik. Chem* **108**, 185 (1924).
- (38) BRÖNSTED A TEETER, *J Phys Chem* **28**, 579 (1924).
- (39) CHRISTIANSEN, *Z. physik. Chem* **113**, 35 (1924).
- (40) CHRISTIANSEN, *Z physik Chem* **117**, 433 (1925).
- (41) DANNER, *J. Am Chem Soc.* **44**, 2824, 2832 (1922)

- (42) DAWSON, J. Chem. Soc. **99**, 1 (1911).
- (43) DAWSON A. CRANN, J. Chem. Soc. **109**, 1272 (1916).
- (44) DAWSON A. POWIS, J. Chem. Soc. **103**, 2135 (1913).
- (45) DAWSON A. REIMANN, J. Chem. Soc. **107**, 1426 (1915).
- (46) DAWSON A. WHEATLEY, J. Chem. Soc. **97**, 2048 (1910).
- (47) DEBYE A. HÜCKEL, Physik Z. **24**, 185 (1923).
- (48) DIMROTH, Ann. **377**, 128 (1910).
- (49) EBERT, Jahrb. Rad. El. **18**, 134 (1921).
- (50) EBERT, Z. Elektrochem. **31**, 113 (1925).
- (51) EULER, Z. physik. Chem. **28**, 619 (1899).
- (52) EULER, Sv. Vet. Akad. Forh. 1899 Nr. 4.
- (53) EULER, Z. physik. Chem. **32**, 348 (1900).
- (54) EULER, Z. physik. Chem. **36**, 641 (1901).
- (55) EULER U. HEDELIUS, Biochem. Z. **107**, 150 (1920).
- (56) EULER, MYRHACK O. RUDBERG, Ark. Kem. Min. Geol. **8**, Nr. 28 (1923),
Z. anorg. allgem. Chem. **127**, 244 (1923).
- (57) EULER O. LAURIN, Ark. Kem. Min. Geol. **7**, Nr. 30 (1920).
- (58) EULER O. SVANBERG, Z. physiol. Chem. **115**, 139 (1921).
- (59) EULER U. ØLANDER, Z. anorg. allgem. Chem. **147**, 295 (1925).
- (60) EULER, ØLANDER U. RUDBERG, Z. anorg. allgem. Chem. **146**, 45 (1925).
- (61) FAJANS, Z. physik. Chem. **73**, 25 (1910).
- (62) FAJANS, Ber. physik. Ges., **21**, 709 (1919).
- (63) FAJANS U. JOOS, Z. physik. **23**, 1 (1924).
- (64) FALES A. MORELL, J. Am. Chem. Soc. **44**, 2071 (1922).
- (65) FAULKNER A. LOWRY, J. Chem. Soc. **127**, 1080 (1925).
- (66) FITZGERALD A. LAPWORTH, J. Chem. Soc. **93**, 2163 (1908).
- (67) FRAENCKEL, Z. physik. Chem. **60**, 202 (1907).
- (68) FRANCIS A. CLIBBEN, J. Chem. Soc. **101**, 2358 (1912).
- (69) FRANCIS A. GEAKE, J. Chem. Soc. **103**, 1722 (1913).
- (70) FRANKLIN, J. Am. Chem. Soc. **46**, 2137 (1924).
- (71) GEFFCKEN, Z. phys. Chem. **49**, 257 (1904).
- (72) GOLDSCHMIDT, Ber. **28**, 3218 (1895).
- (73) GOLDSCHMIDT, Ber. **29**, 2208 (1896).
- (74) GOLDSCHMIDT, Z. Elektrochem. **15**, 4 (1909).
- (75) GOLDSCHMIDT, Z. physik. Chem. **70**, 627 (1910).
- (76) GOLDSCHMIDT, Z. Elektrochem. **17**, 684 (1911).
- (77) GOLDSCHMIDT, Z. physik. Chem. **94**, 233 (1920).
- (78) GOLDSCHMIDT U. BRANAAS, Z. physik. Chem. **96**, 180 (1920).
- (79) GOLDSCHMIDT U. DAHL, Z. physik. Chem. **108**, 121 (1924).
- (80) GOLDSCHMIDT U. DAHL, Z. physik. Chem. **114**, 1 (1925).
- (81) GOLDSCHMIDT, JOHNSEN U. OVERWIEN, Z. physik. Chem. **110**, 251 (1924).
- (82) GOLDSCHMIDT U. THUESEN, Z. physik. Chem. **81**, 30 (1912).
- (83) GOLDSCHMIDT U. UDBY, Z. physik. Chem. **60**, 728 (1907).
- (84) GHOSH, J. Chem. Soc. **113**, 449, 627, 707, 790 (1918).
- (85) GHOSH, J. Chem. Soc. **117**, 823, 1390 (1920).
- (86) GHOSH, Z. physik. Chem. **98**, 211 (1921).
- (87) GRUBE U. SCHMID, Z. physik. Chem. **119**, 19 (1926).
- (88) HALBAN, Z. Elektrochem. **29**, 434 (1923).

- (89) HALBAN, Z. Elektrochem **30**, 601 (1924)
- (90) HANTZSCH, Z. Elektrochem **29**, 221 (1923).
- (91) HANTZSCH, Z. Elektrochem **30**, 194 (1924).
- (92) HARDMAN A. LAPWORTH, J. Chem. Soc. **99**, 2242 (1911).
- (93) HARNED, J. Am. Chem. Soc. **40**, 1461 (1918).
- (94) HARNED A. FLEYSHER, J. Am. Chem. Soc. **47**, 82 (1925).
- (95) HARNED A. PFANSTIEL, J. Am. Chem. Soc. **44**, 2193 (1922).
- (96) HARNED A. SELTZ, J. Am. Chem. Soc. **44**, 1475 (1922).
- (97) HENRY, Z. physik. Chem. **10**, 96 (1892).
- (98) HOLMBERG, Z. physik. Chem. **80**, 587 (1912).
- (99) HUDSON, J. Am. Chem. Soc. **29**, 1572 (1907).
- (100) HÜCKEL, Ergeb. exakt. Naturwiss. **3**, 200 (1924).
- (101) JONES, Z. physik. Chem. **14**, 346 (1894).
- (102) JONES, LAPWORTH A. LINGFORD, J. Chem. Soc. **103**, 252 (1913).
- (103) JONES A. LEWIS, J. Chem. Soc. **117**, 1120 (1920)
- (104) JØRGENSEN, D. Vid. Selsk. Skr., Mathem.-naturv. Afd. (8) **2**, Nr. 1 (1916).
- (105) KAILAN, Monatsh., **27**, 543 (1906).
- (106) KAILAN, Z. physik. Chem. **94**, 111 (1920)
- (107) KARLSSON, Z. anorg. allgem. Chem. **119**, 69 (1921)
- (108) KASTLE, Am. Chem. J. **19**, 894 (1897).
- (109) KENDALL, Proc. Am. Acad. Sc. **7**, 56 (1921)
- (110) KENDALL A. BOOGE, J. Chem. Soc. **127**, 1768 (1925)
- (111) KENDALL, BOOGE A. ANDREWS, J. Am. Chem. Soc. **39**, 2303 (1917)
- (112) KENDALL A. BRAKELEY, J. Am. Chem. Soc. **43**, 1826 (1921)
- (113) KENDALL A. KING, J. Chem. Soc. **127**, 1778 (1925)
- (114) KRÜGER, Z. Elektrochem. **17**, 464 (1911).
- (115) KUHN, Z. physik. Chem. **114**, 44 (1925)
- (116) KUHN U. JACOB, Z. physik. Chem. **113**, 389 (1924).
- (117) KULLGREN, Z. physik. Chem. **37**, 613 (1901)
- (118) VAN LAAR, Arch. Teyler (2) **7**, I, 1 (1900)
- (119) VAN LAAR, Z. anorg. allgem. Chem. **139**, 108 (1924).
- (120) LAPWORTH, J. Chem. Soc. **83**, 598 (1903).
- (121) LAPWORTH, J. Chem. Soc. **85**, 30 (1904)
- (122) LAPWORTH, J. Chem. Soc. **93**, 2187 (1908).
- (123) LAPWORTH, J. Chem. Soc. **107**, 857 (1915).
- (124) LAPWORTH A. PARTINGTON, J. Chem. Soc. **97**, 19 (1910)
- (125) LARSSON, Dissertation, Lund (1924).
- (126) LEWIS, Proc. Am. Acad. Sc. **43**, 259 (1907). Z. physik. Chem. **61**, 129 (1907).
- (127) LEWIS A. RANDALL, J. Am. Chem. Soc. **43**, 1140 (1921).
- (128) LEWIS, A System of Phys. Chem. I, N. Y. Longmans, Green & Co. (1920), p. 423.
- (129) LIVINGSTON A. BRAY, J. Am. Chem. Soc. **45**, 2048 (1923).
- (130) LJUNGGREN, Dissertation, Lund (1925).
- (131) LÖB, Z. Elektrochem. **3**, 42 (1896).
- (132) LOWRY, J. Chem. Soc. **75**, 211 (1899).
- (133) LOWRY, J. Chem. Soc. **83**, 1314 (1903).
- (134) LOWRY, J. Chem. Soc. **127**, 1371 (1925)
- (135) LOWRY A. MAGSON, J. Chem. Soc. **93**, 107 (1908).

- (136) LOWRY A. RICHARDS, *J. Chem. Soc.* **127**, 1385 (1925).
(137) MICHAELIS, *Die Wasserstoffionenkonzentration*, I 2. Aufl. Berlin, Springer (1922), p. 16.
(138) MICHAELIS A. MIZUTANI, *Z. physik. Chem.* **116**, 135 (1925).
(139) MICHAELIS A. RONA, *Biochem. Z.* **49**, 232 (1913).
(140) MILNER, *Phil. Mag.* (6) **23**, 551 (1912), (6) **25**, 743 (1913).
(141) MIZUTANI, *Z. physik. Chem.* **116**, 350 (1925), **118**, 318 (1925).
(142) MIZUTANI, *Z. physik. Chem.* **118**, 327 (1925).
(143) MORAN A. LEWIS, *J. Chem. Soc.* **121**, 1613 (1922).
(144) MORTIMER A. PEARCE, *J. Phys. Chem.* **21**, 275 (1917).
(145) NERNST, *Z. physik. Chem.* **4**, 129 (1889).
(146) NERNST, *Theoretische Chemie* 7 Aufl. Stuttgart, F. Enke (1913), p. 582.
(147) NOYES, *J. Am. Chem. Soc.* **46**, 1080 (1924).
(148) ORTON A. JONES, *J. Chem. Soc.* **95**, 114 (1909).
(149) OSTWALD, *J. prakt. Chem.* (2) **23**, 209 (1881), **28**, 449 (1883), **29**, 385 (1884), **30**, 93 (1884), **31**, 307 (1885).
(150) PALMÆR, *Z. physik. Chem.* **22**, 492 (1897).
(151) POMA, *Medd. Nobelinst.* **2**, Nr. 11 (1912).
(152) RAMBERG, *Syra, Bas och Salt*, Upsala Universitets Årsskrift (1925).
(153) REICHER, *Ann.* **228**, 257 (1885).
(154) RICE, *J. Am. Chem. Soc.* **45**, 2808 (1923).
(155) RIIBER, *Ber.* **55**, 3132 (1922), **56**, 2185 (1923), **57**, 1599 (1924).
(156) RIVETT, *Z. physik. Chem.* **85**, 113 (1913).
(157) RØRDAM, *Studies on Activity*, Københ. (1925).
(158) SCATCHARD, *J. Am. Chem. Soc.* **43**, 2387 (1921).
(159) SCATCHARD, *J. Am. Chem. Soc.* **47**, 684 (1925).
(160) SCATCHARD, *J. Am. Chem. Soc.* **47**, 2098 (1925).
(161) SCHREINER, *Z. anorg. allgem. Chem.* **115**, 181 (1921).
(162) SCHREINER, *Z. physik. Chem.* **111**, 419 (1924).
(163) SENTER, *J. Chem. Soc.* **91**, 460 (1907).
(164) SNETHLAGE, *Z. Elektrochem.* **18**, 539 (1912).
(165) SNETHLAGE, *Z. physik. Chem.* **85**, 252 (1913).
(166) STIEGLITZ, *Rep. Intern. Congr. St. Louis*, **4**, 276 (1904), *Am. Chem. J.* **39**, 29 (1905), **39**, 166 (1905), **39**, 402 (1905), **39**, 437 (1905), **39**, 586 (1905), **39**, 719 (1905), *J. Am. Chem. Soc.* **30**, 1263 (1908), **31**, 1319 (1909), **32**, 221 (1910).
(167) STIEGLITZ, *J. Am. Chem. Soc.* **35**, 1774 (1913).
(168) TAYLOR, *Medd. Nobelinst.* **2**, Nr. 34 (1913).
(169) TAYLOR, *Medd. Nobelinst.* **2**, Nr. 35 (1913), Nr. 37 (1913).
(170) TAYLOR, *Z. Elektrochem.* **20**, 201 (1914).
(171) TAYLOR, *Medd. Nobelinst.* **3**, Nr. 1 (1915).
(172) TAYLOR A. CLOSE, *J. Phys. Chem.* **24**, 1085 (1925).
(173) TAYLOR, *J. Am. Chem. Soc.* **37**, 551 (1915).
(174) THIELE U. LACHMANN, *Ann.* **288**, 267 (1895).
(175) TUBANDT, *Ann.* **339**, 41 (1905), **354**, 259 (1907), **377**, 284 (1910).
(176) URECH, *Ber.* **13**, 1693 (1880), **15**, 2130 (1882), **16**, 762 (1883), **20**, 1836 (1887).
(177) WALDEN, *Z. physik. Chem.* **54**, 228 (1906), **55**, 707 (1906), **94**, 263 (1920).
(178) WARDER, *Ber.* **14**, 1361 (1881).

(179) WEGSCHEIDER, Ber **39**, 1054 (1906)

(180) WILHELMY, Pogg. Ann. **81**, 413, 499 (1850).

(181) WOKER, Die Katalyse (Die chemische Analyse XI—XII) Stuttgart, F. Enke (1910)

APPENDIX (1926–1927)

TO CHAPTER III

During the last two years Euler and Ölander (10, 11, 12, 13) have continued their catalytic studies from the standpoint of the ionization theory of catalysis. The opinion that the theory is supported by the catalytic phenomena observed in mutarotation seems also to be maintained in these recent papers, but a theoretical proof is still lacking. In order to explain the phenomena in the catalysis of mutarotation by acid and basic molecules other than hydrogen and hydroxyl ions Ölander (21), however, seems prepared to admit that a compound formed by the substrate and the catalysts in question (e.g., an acetate ion or a pyridine molecule) may well be taken as the reactive molecule characteristic of the theory. He is obviously unaware of the fact that when abandoning the idea of simple substrate ions as the reactive molecules in catalysis and permitting any complex, formed by the substrate and a catalyst, to perform the functions which have hitherto been attached to the substrate ions he actually gives up the fundamental principles of the Euler theory and in substance subscribes to the extended theory advocated in the present review.

A Skrabal (22) has subjected the theory of Euler to an elaborate criticism in an article in which he also gives a review of his own work on hydrolytic reactions from the standpoint of structural chemistry. Strangely enough he seems to believe in the possibility of distinguishing between the two different mechanisms of catalysis by extending experiments to such concentrations of the catalyzing hydrogen and hydroxyl ions that a considerable part of the substrate is transformed into an ion. It is not clear, however, how this conclusion has been arrived at. If it were true, the quite recent paper by Euler and Ölander (13) dealing with the hydrolysis of acetamide by strong acids would have brought good evidence in favor of the ionization theory, as actually believed by these authors. However, it should be emphasized once more that the above disproof of the possibility of discriminating between the two pictures is general and conclusive, so it cannot be avoided by any special construction as to the nature and extent of the reaction considered.

TO CHAPTER IV

The question of primary kinetic salt effect has been dealt with by F. G. Soper (23), who derived a velocity expression for ionic reactions similar to Expression (4, IV). His formula deviates from (4), however, in that the exponent is two-thirds of the exponent in this equation. This divergence is due to a confusion of total and differential free energy (3, 5, 15). Furthermore it appears from another article by the same author (24) that the thermodynamic ideas underlying the derivation of the proposed formula are absolutely unsound, particularly as regards the free energy of the solvent and the part played by it in the charging processes of the Debye-Hückel theory.

Attempts to sustain different forms of the activity rate theory for the case of concentrated salt solutions have been published recently by Åkerlöf (1) and by Soper and Pryde (25). While Åkerlöf from measurements of the rate of decomposition of diacetone alcohol in alkaline solution concludes that the velocity is proportional to the activity of the hydroxyl and the alkali metal ion, Soper and Pryde, studying the rearrangement of chloroacetanilide catalyzed by hydrochloric acid—that is by H^+ and Cl^- —find the speed proportional to the activity of *all* three molecules reacting. From the standpoint of the general activity rate theory the two papers are incompatible. The last mentioned authors in arguing against Equation (2, IV) as applicable to their reaction in concentration up to 1.5 N HCl give no consideration to the fact that this equation is derived only for the domain of much more dilute salt solutions.

L. E. Bowe (4) has studied cane sugar inversion and ester hydrolysis by hydrochloric acid in the presence of concentrated alkali halides and has also determined electrometrically the hydrogen ion activity (called by him the apparent H^+) in the same solution. While for the cane sugar inversion a certain parallelism between speed and hydrogen ion activity was demonstrated, the ester hydrolysis was found to be much less affected by addition of the neutral salt. As was to be expected no general kinetic conclusion could, therefore, be arrived at on the basis of these experiments. Theoretically the work is characterized by the serious mistake of confusing activity and osmotic pressure and by making use of the concept of solution pressure, which, except for solutions following the gas laws, has no definite significance. In this respect he seems to be a follower of W. Bancroft (2) who evidently has not yet familiarized himself with the modern conceptions in the field of solutions.

TO CHAPTER V

M. Kilpatrick (18) has studied catalysis in buffer solutions and shown the presence of a secondary kinetic salt effect. Otherwise the importance of the theory of primary and secondary salt effect does not appear to be recognized as yet, even in very recent work on catalysis. Thus Dawson and Carter (8) still believe in the possibility of calculating hydrogen ion concentrations by the classical mass action expression and the simple Nernst formula in salt solutions the concentration of which varies from 0 to 1 normal. It will appear from the above theory of salt effect that the error involved in this way of calculation may easily amount to 100 per cent or more. Lowry and Smith (20) in their paper on mutarotation of glucose attribute high importance to Dawson's work on account of its "repudiation" of the theory of secondary salt effect as first demonstrated in Arrhenius' experiments quoted above. They fail to realise, however, that the effect of sodium chloride in accelerating the inversion of cane sugar by a weak acid is entirely different in nature from the accelerating effect of sodium acetate in the mutarotation of glucose. In the interpretation of catalytic phenomena in salt solutions there is no escape from the obligatory condition of considering three entirely different effects: 1, primary salt effect, 2, secondary salt effect, 3, direct catalytic effect of the ions in conformity with the extended theory of acid and basic catalysis presented in chapter VIII of the present review. Owing to the magnitude of the last mentioned influence in the acetone-iodine reaction the conclusions of Dawson and Carter for this reaction are qualitatively correct in spite of their neglecting the effects mentioned under 1 and 2. On the other hand in the investigation by Dawson and Lowson (9) on ester hydrolysis the same omission becomes fatal.

TO CHAPTER VI

The general question of the definition of an acid has been dealt with by A. F. Germann (14) and A. Hantzsch (17). The ideas of the first of these authors, although of a subversive tendency, are extremely vague and more likely to confuse than to clarify the conception. Hantzsch describes an acid as a compound of hydrogen with a negative group. Beside the uncertainty implied by using the undefined idea of a "negative" group the description of Hantzsch is not actually a definition, but rather an attempt to point out which characteristics are required to make a substance an acid. The inadequacy of the acid

"dissociation" as a general measure of acidity is obviously what makes Hantzsch fall back on a more or less obsolete standpoint, as expressed by his considering "salt formation" as a property of fundamental significance for the concept of acidity. This idea disagrees entirely with our point of view. However, the inadequacy of measuring acidity in different solvents by means of dissociation has also been emphasized above, and we may find after all—in spite of much divergency—that the ideas of Hantzsch are, in many ways, not incompatible with those of the present review.

Conant and Hall (7, 16) have made an interesting investigation on acidity and basicity in acetic acid as a solvent and have shown how the possibility of titrating an acid or a base depends upon the acid and basic properties of the solvent.

TO CHAPTER VIII

The investigation on mutarotation referred to above has now been published (6). The results obtained are in full conformity with the extended theory of catalysis showing a marked catalytic effect of a number of acid and basic molecules. In a logarithmic plot of catalytic constant against strength constant for all the different electric types of catalysts including the solvent the results come out as a straight line with a slope of 0.4 and about 0.2 for bases and acids respectively. This relationship seems approximately fulfilled throughout a range of acid and basic strength of about 10^{18} , and it is possible therefore within this range to predict approximately the catalytic efficiency of any base or acid in the mutarotation merely from its strength.

Lowry and Smith (20) have studied the glucose reaction from a similar point of view and have obtained values for acetic acid, acetate and ammonium ion which agree well with our results. The value for the catalytic constant of the hydrochloric acid molecule cannot be accepted, however, because the authors have based their calculations upon the classical view of dissociation of strong electrolytes. Some aspects developed by Lowry (19) in a subsequent paper on the mechanism of catalysis are much in accord with the extended theory of catalysis presented above.

The paper by Dawson and Carter (8) on the acetone-iodine reaction mentioned above, is a continuation of previous work on the basis of the dual theory by Dawson and his coworkers. The new work, however, contains much more conclusive evidence in favor of the extended

theory of catalysis although, as mentioned, the omission of salt effect from their considerations makes quantitative calculation difficult.

It should be mentioned finally that experiments now in progress in this Laboratory have brought further support to the ideas of the extended theory of acid and basic catalysis.

LITERATURE TO APPENDIX

- (1) ÅKERLOF, J Am Chem. Soc. **49**, 2955 (1927).
- (2) BANCROFT, J Phys Chem **30**, 1194 (1926).
- (3) BJERRUM, Z. physik Chem. **119**, 145 (1926).
- (4) BOWE, J Phys. Chem **31**, 291 (1927).
- (5) BRØNSTED, Trans. Faraday Soc. **23**, 416 (1927)
- (6) BRØNSTED AND GUGGENHEIM, J Am. Chem Soc. **49**, 2554 (1927).
- (7) CONANT AND HALL, J Am. Chem. Soc. **49**, 3062 (1927).
- (8) DAWSON AND CARTER, J. Chem. Soc. 2282 (1926).
- (9) DAWSON AND LOWSON, J Chem Soc. 2444 (1927).
- (10) EULER AND ÖLANDER, Z anorg allgem Chem **152**, 113 (1926).
- (11) EULER AND ÖLANDER, Z anorg allgem Chem **156**, 143 (1926).
- (12) EULER AND ÖLANDER, Z Elektrochem **33**, 527 (1927).
- (13) EULER AND ÖLANDER, Z. physik Chem **131**, 107 (1927).
- (14) GERMANN, J Am Chem. Soc **47**, 2461 (1925)
- (15) GÜNTELBURG, Z physik Chem **123**, 199 (1926)
- (16) HALL AND CONANT, J Am Chem. Soc **49**, 3047 (1927)
- (17) HANTZSCH, Ber **60**, 1933 (1927)
- (18) KILPATRICK, J Am Chem Soc **48**, 2091 (1926)
- (19) LOWRY, J Chem Soc. 2554 (1927).
- (20) LOWRY AND SMITH, J. Chem. Soc 2539 (1927).
- (21) ÖLANDER, Z physik. Chem. **129**, 1 (1927).
- (22) SKRABAL, Z Elektrochem. **33**, 322 (1927).
- (23) SOPER, J. Phys Chem **31**, 1790 (1927).
- (24) SOPER, J. Phys Chem. **32**, 67 (1928).
- (25) SOPER AND PRYDE, J Chem. Soc. 2761 (1927).

FREE INORGANIC RADICALS¹

P. WALDEN² AND L. F. AUDRIETH³

Baker Laboratory of Chemistry, Ithaca, N. Y.

INTRODUCTION

A radical has been defined as a group of elements which, in reactions and properties, behaves as an entity. The study of *organic* radicals, beginning with the classical researches of Gomberg, Schlenk, Wieland and others (1), has long since led to many interesting results. However, the study of *inorganic* radicals, in particular of negative univalent inorganic groups, is one of the most recent developments in pure chemical science. Because of the similarity of certain of these groups to the halogens, they have been termed "halogenoids" by Browne and his co-workers (2), and "pseudo-halogens" by Birckenbach and Kellermann (3). At present we recognize as "halogenoid" radicals the thiocyanate, cyanide, azidodithiocarbonate, selenocyanate, azide, cyanate, fulminate, and tellurocyanate groups, of which the first four have been isolated in the free dimolecular form.

I. HISTORICAL

Cyanogen, $(\text{CN})_2$, was first obtained by Gay-Lussac in 1815 by heating mercuric or argentous cyanide. From the standpoint of theoretical chemistry this discovery was very important, as it is the first recorded instance of the isolation of a "radical" which could properly be classed under the definition given in 1838 by

¹ The term "free radical" is not used here in the same sense that the organic chemist speaks of "free organic radicals" (Ref. (1).) The latter, as non-saturated substances, are capable of existing in the mono-molecular state, such as triphenyl methyl, $(\text{C}_6\text{H}_5)_3\text{C}$, whereas the inorganic radicals have been obtained hitherto only in the dimeric form—just as chlorine, Cl_2 , is spoken of as "free chlorine."

² George Fisher Baker Non-Resident Lecturer in Chemistry at Cornell University.

³ Heckscher Research Fellow in Chemistry at Cornell University.

Liebig (4), who said, "We call cyanogen a radical, (a) because it does not change its constitution in a series of compounds; (b) because in these compounds it may be replaced by simple substances (elements); (c) because in its compounds with an element the latter may be liberated and replaced by its equivalent of another element."

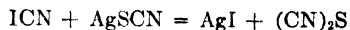
Soon after the discovery of cyanogen Liebig (5) attempted to prepare free *rhodan* or thiocyanogen by passing chlorine over dry silver or lead thiocyanate. In both cases he observed a strong evolution of heat; the compound became yellow, then increased in volume, became more viscous and finally evolved a red vapor which condensed to a red solid. The whole mass eventually hardened and became vermillion in color. The presence of sulfur chloride and cyanuric chloride in the colder portions of the tube seemed to indicate a decomposition of the thiocyanate radical. No product was found which gave even approximate analyses for thiocyanogen. Consequently Liebig concluded that the red product obtained by him was composed of a mixture of substances.

Liebig next passed chlorine through a hot concentrated solution of potassium thiocyanate and obtained a yellow amorphous precipitate, which approximated the composition of thiocyanogen. It was found to be insoluble in all common solvents, soluble, however, in concentrated sulfuric acid. In announcing the discovery of the supposed thiocyanogen, he mentioned the fact that its inert nature and non-volatility were contrary to expectations.

Other investigators worked with the same substance and found its composition to vary considerably. It was ascertained, furthermore, that the substance contained hydrogen and oxygen in addition to sulfur, carbon, and nitrogen. Liebig later (1844) admitted that the substance was probably a mixture of various compounds.

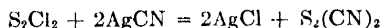
Subsequently it was found that the action of such oxidizing agents as potassium permanganate, nitric acid, hydrogen peroxide, and potassium persulfate on thiocyanates would also give similar yellow products. These were then collectively given the name "pseudothiocyanogen" (and still later "canarine").

In 1861 Linnemann (6) attempted to prepare the thioanhydride of thiocyanic acid, probably not knowing that Lassaigne (1828) had already obtained cyanogen sulfide by the interaction of sulfur chloride and mercuric cyanide. He succeeded in preparing the same substance from cyanogen iodide and silver thiocyanate:



He next allowed silver thiocyanate to react directly with an ethereal solution of iodine and obtained silver iodide and a readily volatile red-brown liquid. Because of the instability of the latter he did not pursue his investigations any further. On permitting the solution to stand for a time, large quantities of a yellow amorphous substance formed, which led him to conclude that the freshly prepared solution contained iodine thiocyanate, a statement which has never yet been denied or affirmed by any other investigator.

Schneider (7) allowed sulfur chloride in carbon disulfide solution to react with silver cyanide expecting to get a transformation in accordance with the equation:



A violent reaction took place and "white crystals with a penetrating odor invoking tears" separated out, possessing the theoretical composition $(\text{CN})_2\text{S}_2$. Microscopic study revealed the fact that the product was heterogeneous, and the assumption of the formation of the disulfide was given up. Later the substance was recognized to be a mixture of "dicyanogen mono-sulfide and dicyanogen tri-sulfide" (8).

Klason (9) oxidized thiocyanuric acid with iodine and obtained a compound to which he assigned the formula $(\text{CN})_3\text{S}_3 \cdot \text{S}_3(\text{CN})_3$, an empirical polymer of thiocyanogen. In summarizing the work which had been done up to 1901 Goldberg (10) came to the conclusion that free thiocyanogen had not yet been prepared.

Negative results of this nature seemed to make it only too evident that radicals could not be liberated from their compounds and that they existed as entities only in solution, that is, in the

ionic state. These views were expressed by Wm. Ostwald in 1895, who stated "Radicals by their inherent nature are not capable of isolation".

Thus, it was not until 1920 that thiocyanogen, the first halogenoid, was obtained by Söderbäck (13), although Sommer (11) had previously obtained azido-carbondisulfide, $(\text{SCSN}_3)_2$, without recognizing its true nature or studying it. It is the work of this brilliant Swedish investigator which gave a new impetus to the study of "inorganic radicals," and it is the appearance of his article dealing with this subject which marks the beginning of renewed interest in their characteristics and properties.

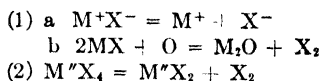
II. PROPERTIES OF THE HALOGENS

In order to bring out the halogen-like character of these substances, which we have been wont to call "halogenoids", it may be of value to discuss briefly those properties which characterize the halogens as a *group*. The great similarity existing among the latter, both as elements and in their compounds, is such that the properties of any one member can be said to typify those of any other member, even though fluorine shows a little deviation from some of these.

- (1) The halogens, in general, show pronounced volatility.
- (2) The solid substances are apparently isomorphous, belonging to the rhombic system.
- (3) They show great affinity towards metals, combining directly in many cases to form salts. Their silver, mercurous, and lead salts are to a varying degree insoluble in water. They are, however, soluble in potassium cyanide and sodium thiosulfate solutions.
- (4) With hydrogen they form hydracids, compounds which in aqueous solution are usually highly ionized.
- (5) They are capable of forming inter-halogen compounds, such as ClI , ClI_3 , BrCl , etc.
- (6) They form polyhalide complexes, such as CsCl_2I , $\text{CsI} \cdot \text{I}_4$.
- (7) Oxyacids have been prepared of which the most notable series is that of chlorine.

(8) They form characteristic double and complex salts, such as $\text{CuCl}_2 \cdot 2\text{KCl}$ and K_2PtCl_6 , respectively.

(9) The free halogens may be prepared in general, (1) by chemical or electrochemical oxidation of the hydric acids or their salts, or (2) by decomposition of the perhalides.



III. THE HALOGENOIDS

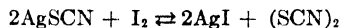
Let us apply these criteria to those substances which we have termed "halogenoids." This name has been used to designate any univalent chemical aggregate composed of two or more electronegative atoms, which shows in the free state certain characteristics of the halogens, and which combines with hydrogen to form an acid, and with silver to form a salt insoluble in water. The halogenoid salts and hydric acids have been known to science for many years. The silver, lead, and mercurous salts are insoluble in water. These few similarities could hardly be considered sufficient basis for further generalizations. However, the fact that the radicals, in so far as they have been obtained in the free state, do simulate the halogens in behavior, makes such an analogy permissible.

As determined from a study of the electrical conductivity of some of their salts in aqueous and alcoholic solutions, halogen and halogenoid ions have been arranged in the following activity or electromotive series (3), viz.: F^- , ONC^- , OCN^- , Cl^- , N_3^- , Br^- , CN^- , SCN^- , SCSN_3^- , I^- , SeCN^- , TeCN^- . It is of interest to note the explanation which Birkenbach and Kellermann offer in accounting for the halogenoid properties of these substances. On the basis of the octet theory if one considers the sum of the valence electrons of the atoms comprising the halogenoid group, rather than the individual valence electrons, it is possible to formulate an arrangement of saturated shells of eight electrons (octets), and an additional structure of seven electrons, similar to that of the halogen atoms. For example, the azide group will have a total of fifteen ($8 + 7$); the selenocyanate again a

total of fifteen ($8 + 7$); the azidodithiocarbonate group 31 ($8 + 8 + 8 + 7$). The exact configuration of the electrons is not made plain. Although such an explanation might be extended to account for the observed similarities it would unquestionably lead to wrong conclusions, *e.g.* the nitrate radical which has a total of twenty-three electrons ($23 = 2 \times 8 + 7$), and the hydroxyl radical which has a total of seven. Darrow's (12) statement might properly be cited here: "Undoubtedly the most prudent policy is to consider atoms as being creations of the scientific imagination, mental mechanisms designed to copy particular restricted sets of phenomena."

To bring out the halogenoid analogy, let us discuss: methods of preparation, physical and chemical properties of those which have been isolated in the free state; formation of compounds with the halogens (interhalogen-halogenoids) and among themselves (interhalogenoids); compounds corresponding to the polyhalides.

I. *Thiocyanogen*, $(SCN)_2$. Free thiocyanogen was first obtained by Söderbäck (13) who allowed iodine to react with an ethereal suspension of silver thiocyanate. Various solvents were tried; bromine was used in place of iodine, and various other metallic thiocyanates in place of the silver salt. Bromine was found to be much more satisfactory than iodine, as reaction with the latter comes to an equilibrium and does not go to completion.

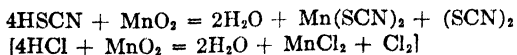


The velocity of the reaction is markedly influenced by the solvent employed; thus a N/10 bromine solution in carbon disulfide is decolorized in a few minutes by lead thiocyanate, while a considerably longer period of time is necessary when bromine is dissolved in ether. Other factors also enter into the reaction. If the carbon disulfide is dried over phosphorous pentoxide, the bromine is not decolorized by the lead salt even after an hour, although the reaction is immediate when silver or mercuric thiocyanates are used. Solutions of thiocyanogen so obtained liberate iodine from iodides, oxidize copper from the cuprous to the cupric state, and combine directly with metals to form the corresponding thiocyanates.

Solid thiocyanogen is prepared by evaporating a $N/2$ solution in ethyl chloride, ethyl bromide, or ether, over sulfuric acid, and cooling the concentrated solution in a special filtering apparatus to -70° . The thiocyanogen is precipitated out, filtered, and dried at this temperature and is obtained as a crystalline substance composed of "cruciform aggregates" which melts at -2 to -3° to a viscous yellow oil which may be supercooled to -20°C . It always solidifies at about -30° to small rhombic plates. If liquid thiocyanogen is warmed to room temperatures it gradually becomes less viscous and changes in color to a dark red-brown. At a certain stage a yellow vapor is given off with the evolution of a good deal of heat, and the material solidifies to a brick-red, amorphous mass. Analysis seems to indicate that it is fairly pure polymeric thiocyanogen or "parathiocyanogen," a product which is insoluble in all ordinary solvents and has, thus far, defied definite experimental treatment. Potassium cyanide converts it into a black compound, which is soluble in alkalis. Analysis of the black compound indicates the formula $(\text{CN})_2\text{SOH}$.

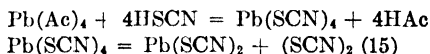
Previous to its isolation, the existence of thiocyanogen as a product of decomposition of gold thiocyanates was demonstrated by Bjerrum and Kirschner (14). By electrometric measurements these investigators determined it to be an unstable halogen-"oid," intermediate between bromine and iodine. They found that bromine was instantly reduced by an acidified solution of sodium thiocyanate, and that the resulting colorless solution set free iodine from potassium iodide.

Solutions of thiocyanogen may also be prepared by oxidation of the free acid in ethereal solution by manganese dioxide (15). This method is not suggested as feasible since the yields are extremely small. The reaction is of interest, however, since it parallels the one which is used in the laboratory preparation of chlorine:



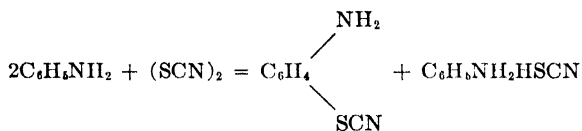
Like the halogens, thiocyanogen may also be prepared by electrolysis of thiocyanates, preferably in alcoholic solution (16). Inter-

action of lead tetracetate and thiocyanic acid in ethereal solution presumably takes place in accordance with the following equations:



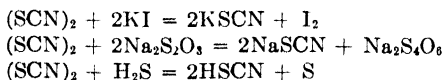
The molecular weight of thiocyanogen in bromoform has been determined by Lecher and Goebel (17). Cryoscopic measurements for solutions less than half-normal indicate it to be dissolved as $(\text{SCN})_2$. In solutions greater than half-normal there are evidences of polymerization.

Many reactions indicate the similarity of thiocyanogen with the halogens. In addition to those mentioned previously it has been found that thiocyanogen reacts with mercury diphenyl to yield phenyl mercuric thiocyanate and with aniline to give p-thiocyananiline and aniline thiocyanate:



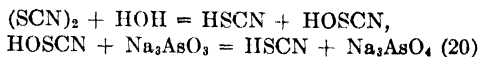
The action of ethyl nitrite upon thiocyanic acid yields a nitrosyl thiocyanate (analogous to NOCl), which is unstable above -17°C ., decomposing above that temperature into nitric oxide and thiocyanogen (18).

Kaufmann and Gaertner (19) have worked out a series of thiocyanometric determinations, which make it possible to titrate solutions of thiocyanogen in organic solvents with aqueous solutions of potassium iodide, hydrogen sulfide, and sodium thiosulfate.

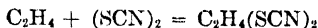


In the presence of a sufficient quantity of salts, hydrolysis of thiocyanogen is reduced to such an extent that the above reactions take place. The first reaction is most suitable for the quantitative determination of thiocyanogen in organic solvents.

Sodium arsenite is oxidized by thiocyanogen to sodium arsenate. The intermediate formation of hypothiocyanous acid is postulated:



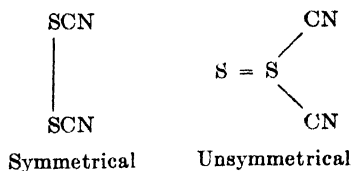
Like the halogens, thiocyanogen reacts directly with unsaturated hydrocarbons to form addition products (21):



From their work on the gold thiocyanates, Bjerrum and Kirschner conclude that the hydrolysis of thiocyanogen proceeds in accordance with the equations:

- (1) $(\text{SCN})_2 + \text{HOH} = \text{H}^+ + \text{SCN}^- + \text{HOSCN}$ (instantaneous)
- (2) $2\text{HOSCN} = \text{HSCNO}_2 + \text{H}^+ + \text{SCN}^-$ (rapid)
- (3) $2\text{HSCNO}_2 = \text{HSCNO} + \text{HSCNO}_3$ (instantaneous)
- (4) $\text{HSCNO}_3 + \text{HOH} = \text{H}_2\text{SO}_4 + \text{HCN}$ (instantaneous)
- (5) $3(\text{SCN})_2 + 4\text{H}_2\text{O} = \text{H}_2\text{SO}_4 + \text{HCN} + 5\text{HSCN}$

However, Lecher and his co-workers have recently found that the reaction does not proceed quantitatively in accordance with equation (5). In view of the fact that the polymerization of thiocyanogen has been found to take place with formation of small quantities of cyanogen thiocyanate and dithiocyanogen mono-sulfide (22), it has been assumed that the halogenoid is capable of existing in two tautomeric forms, viz;



Based on analogy with sulfur chloride, $\text{S} = \text{S} \begin{array}{l} \diagup \text{Cl} \\ \diagdown \text{Cl} \end{array}$, unsymmetrical

thiocyanogen with the disulfide structure might be expected to *hydrolyze* in accordance with the equations:

- (1) $2(\text{S}=\text{S}(\text{CN})_2) + 4\text{HOH} = 4\text{HCN} + 2\text{H}_2\text{S} + 2\text{SO}_2$
- (2) $\text{SO}_2 + 2\text{H}_2\text{S} = 2\text{H}_2\text{O} + 3\text{S}$



Equation (1) is based upon analogy to the sulfur halides, (2) and (3) are well known, while (4) is indicated from the reaction between sodium arsenite and thiocyanogen. The sum total of all of the above reactions would lead to the same final result as postulated by Soderbäck, Bjerrum, and others, although under the circumstances it is quite probable that other complications arise which do not permit of definite quantitative expression.

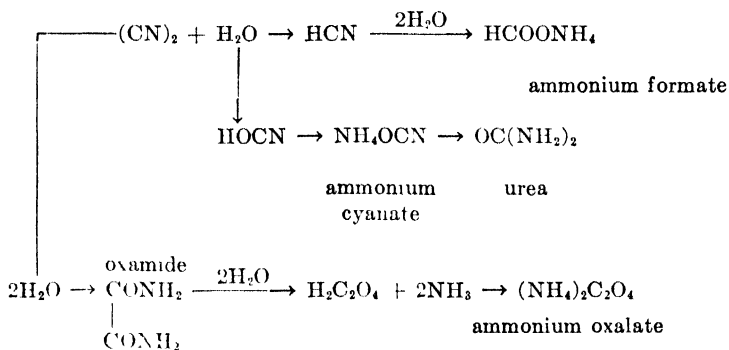
Aminolysis of thiocyanogen results in formation of compounds analogous to the chloramines:



With ammonia, thiocyanamide, $SCN \cdot NH_2$, an extremely unstable oil is obtained (23).

II. *Cyanogen*, $(CN)_2$. Cyanogen is an extremely poisonous gas, colorless and possessing a peculiar odor. It is liquefiable at ordinary pressures at $-25^\circ C$. When heated to $500^\circ C$., some of the gas is converted into the insoluble polymer, paracyanogen, $(CN)_x$. Its reactions have been common chemical knowledge for over a century—and its halogenoid characteristics have been the subject of considerable investigation, so that a detailed discussion of its various properties here would be quite superfluous. However, two interesting analogies may be brought out.

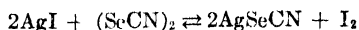
It undergoes *hydrolysis*, the course of which may be represented diagrammatically:



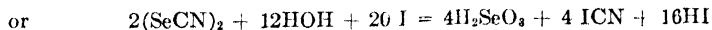
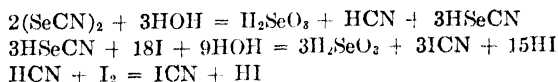
Ammonolysis results in formation of cyanamide and ammonium cyanide.

III. *Selenocyanogen*, $(\text{SeCN})_2$. Selenocyanogen is obtained as a homogeneous yellow powder, which within a very short time turns red in color. If kept dry and in a vacuum it is stable for a longer period. It is soluble in benzene, chloroform, and carbon tetrachloride. Solutions in acetic acid decompose rapidly with precipitation of selenium. Molecular weight determinations in benzene are in fair agreement with the dimeric formula $(\text{SeCN})_2$. Determinations in glacial acetic acid would seem to indicate a somewhat marked dissociation—possibly with formation of the unsaturated radical. This apparently anomalous behavior is explained by chemical interaction of solvent and solute.

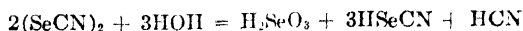
Selenocyanogen may be prepared by electrolysis of a solution of potassium selenocyanate in methyl alcohol (3), or by the decomposition of lead tetraselenocyanate (24), as formed by the interaction of selenocyanic acid and lead tetracetate in acetone. The reaction between iodine and silver selenocyanate has also been studied, and, both electrically and analytically, has been found to proceed in both directions to an equilibrium, where 86 per cent of the iodine is converted to silver iodide, and the remainder left in ethereal solution:



The strength of selenocyanogen solutions may be determined by titration with aqueous iodine. The following reactions are assumed to take place:



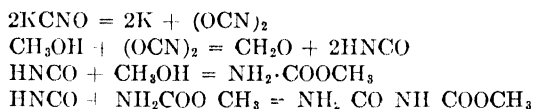
Upon *hydrolysis*, selenocyanogen yields selenous, hydrocyanic, and selenocyanic acids, presumably in accordance with the following equation:



In view of the fact that dicyanogen selenide and triselenide are formed on heating selenocyanogen in carbon disulfide, it is assumed that the latter may also, like thiocyanogen, exist in two tautomeric forms:

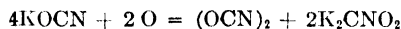


IV. *Oxycyanogen*, $(\text{OCN})_2$. The presence of this halogenoid in solution has been demonstrated by Birckenbach and Kellermann (3) who attempted to prepare it by the electrolysis of a methyl alcoholic solution of potassium cyanate. Instead of the free halogenoid they obtained methyl allophanate and detected the presence of formaldehyde, indicating that the following reactions may have taken place:



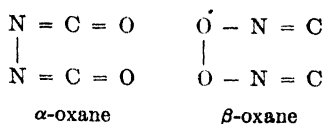
However, the presence of some free oxycyanogen was demonstrated, as the solutions obtained liberated iodine from potassium iodide, and dissolved copper, zinc and iron without gas evolution.

Lidov (25) claims to have prepared oxycyanogen or "oxane" (a) by the action of hydrogen peroxide, cupric oxide, or sodium hypobromite upon potassium cyanate:



(b) by reduction of nitrogen dioxide by carbon at 150° , and (c) by the interaction of cyanogen bromide and silver oxide. According to Lidov, oxane is a gas similar to carbon dioxide, and may be present in human exhalations. It is supposed to be the anhydride of oxanic acid, H_2CNO_2 , of which several salts have been described.

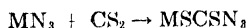
By varying the conditions of procedure, two isomeric oxanes may be produced:



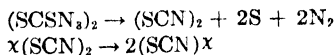
Apparently this work has either been overlooked by recent investigators, or its reliability questioned. There is no doubt but that substantiation of the above observations would be desirable.

V. *Azidocarbondisulfide*, $(\text{SCSN}_3)_2$. This substance was first studied by Browne and his co-workers (2), who prepared it by chemical and electrochemical oxidation of azidodithiocarbonic acid and its salts. It is practically insoluble in water and can therefore easily be obtained by addition of an iodine-potassium iodide solution to an aqueous azidodithiocarbonate solution. It is a white crystalline solid, very unstable, and sensitive to both shock and impact.

It is of particular interest as its preparation and decomposition involve the use and appearance of two other halogenoid radicals. The azidodithiocarbonates (26) are prepared by interaction of carbon disulfide and azides:



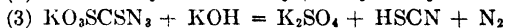
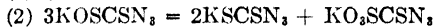
and the free halogenoid itself decomposes autocatalytically to give thiocyanogen and eventually its polymer:



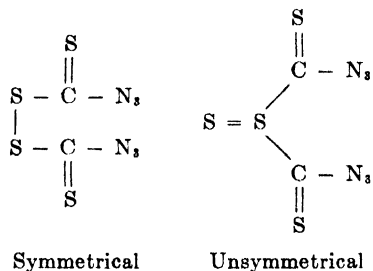
The acceleration in the rate of decomposition of the halogenoid, as noted during the course of nitrometric studies, is thought to be due to the intermediate formation of an unstable *interhalogenoid* complex, possibly $\text{SCN} \cdot \text{SCSN}_3$.

The halogenoid dissolves in alkali with formation of a greenish solution. Primary *hydrolysis* is assumed to take place in accordance with the equations (1), (2), and (3), in which the transient

formation of the hypochlorite and chlorate analogues of $(\text{SCSN}_3)_2$ are postulated.



That this reaction does not hold quantitatively has been shown by Gardner (27), who has found thiosulfates, sulfites, and sulfides among the products. Here also, it is assumed that azidocarbonyldisulfide may exist in two tautomeric forms:



Ammonolysis takes place with formation of ammonium azido-dithiocarbonate (28) and an oily, unstable product, presumably the amide derivative, NH_2SCSN_3 :



In connection with azidocarbonyldisulfide, it might be of interest to mention that thiuram disulfide $(\text{SCSNH}_2)_2$ and its substitution products, $(\text{SCSNHR})_2$ and $(\text{SCSNR}_2)_2$ are all structurally related to the halogenoid under discussion. Certain reactions and properties, such as the preparation of the dimeric radical and of the corresponding acids and salts, would make it advisable to study these substances more closely from the standpoint of the halogenoid analogy. Enough evidence is now at hand to warrant their designation as "organic halogenoids."

VI. *Other halogenoids.* Attempts have been made to prepare the other halogenoid radicals in the free form. Thus, Browne and Lundell (29) electrolyzed a solution of potassium azide in anhydrous liquid hydrogen azide, but obtained only molecular

nitrogen. *Tellurocyanogen*, $(\text{TeCN})_2$, has thus far been investigated only very superficially. It should be capable of isolation, either by the electrolysis of solutions of its potassium salt in methyl alcohol, or by the interaction of potassium tellurocyanate and lead tetracetate in acetone solution.

The free radical corresponding to the *fulminate* ion has thus far not been subjected to experimental treatment. Sufficient reasons appear to be at hand, however, to permit its classification with the halogenoid groups. It may certainly be predicted that all of these substances will some day be obtainable as "free radicals."

IV. POLYHALOGENOIDS

Compounds of the type $\text{MX} \cdot \text{X}_2$ (where M = alkali metal and X = halogen) have become firmly established and recognized as chemical individuals (30). One of the most definite supports of the halogenoid analogy is the isolation of several polyhalogenoids and evidence of the existence of others.

A compound intermediate between the polyhalides and the polyhalogenoids is cesium diiodocyanide, CsI_2CN (31), prepared by the interaction of cesium iodide and cyanogen iodide in aqueous solution as a reddish crystalline compound. The corresponding potassium compound is also known, but is much less stable.

Kerstein and Hoffman (16) have shown that complex compounds of the trithiocyanate type are possible. Ammonium trithiocyanate, $\text{NH}_4(\text{SCN})_3$, has been isolated as a clear crystalline substance, stable below -6° . The excess thiocyanogen can be removed by treatment with carbon disulfide. Electrolysis of a cold concentrated aqueous solution of potassium thiocyanate results in formation of free thiocyanogen, which is supposedly present in the form of the polyhalogenoid, $\text{KSCN} \cdot (\text{SCN})_2$. Experiment has shown that hydrolysis of thiocyanogen is retarded by the presence of a large excess of an alkali thiocyanate—presumably because of the formation of such polyhalogenoids.

Selenocyanogen is markedly more stable in solutions of potassium selenocyanate. Again it is supposed that compounds cor-

responding to polyhalides are formed, and this seems to be indicated by the work of Birckenbach and Kellermann (32), who, by electrometric titration of solutions of potassium selenocyanate and of potassium iodide with iodine, have established the existence of a series of complex salts, *e.g.* KI_3 , $\text{K}(\text{SeCN})\text{I}_2$, $\text{K}(\text{SeCN})_2\text{I}$, and $\text{K}(\text{SeCN})_3$. These are all unstable in presence of water.

Verneuil (33) claims to have prepared a compound possessing the formula $\text{K}(\text{SeCN})_3$, stable in water. It is rather questionable whether this is at all similar to the cesium tri-selenocyanate prepared by Birckenbach and Kellermann. The latter compound is a red-brown crystalline substance, stable in air, but decomposed immediately by water.

V. INTERHALOGEN-HALOGENOIDS

From the list of interhalogen compounds theoretically capable of existence as definite chemical individuals, bromine trifluoride (34), iodine pentafluoride (35), iodine monochloride and trichloride (36) and iodine monobromide (37) have been prepared.

Like their prototypes, the halogenoids are capable of entering into combination (*a*) with the halogens to form interhalogen-halogenoid compounds, and (*b*) with other halogenoids to form interhalogenoids.

The interhalogen-halogenoids include chlorazide, ClN_3 (38), bromazide, BrN_3 (39), iodoazide, IN_3 (40), the chloride, bromide, and iodide of cyanogen (all of which have been long recognized as chemical individuals), and thiocyanogen monochloride and trichloride (41).

The halides of "azine," $(\text{N}_3)_2$, are all extremely explosive substances, which undergo spontaneous decomposition at room temperatures. In so far as they have been studied they resemble the cyanogen halides very closely. They are characterized by their intense odors, and have the physiological effects of hydrazoic acid.

The cyanogen halides are very volatile substances whose vapors are extremely poisonous and possess a very pungent odor. They have long been known and have found wide application in synthetic organic chemistry.

There has been some question as to the molecular weight of thiocyanogen chloride, SCNCl . One might expect this substance to be an extremely volatile liquid, but, as obtained by direct combination of thiocyanogen and chlorine in chloroform, it is a white crystalline solid, non-volatile, decomposing only over 200°C . Lecher and Joseph repeated the original experiments and found the substance to possess a molecular weight corresponding to $(\text{SCNCl})_6$ in benzene. The trichloride is obtained by direct combination of chlorine and thiocyanogen in ethyl bromide, as a yellow oil, with a pungent odor, boiling at $152\text{--}153^\circ$.

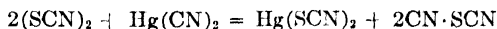
Browne and Gardner (42) have recently studied the action of the halogens on azidocarbonyldisulfide and have obtained qualitative evidence of the existence of $\text{Cl}\cdot\text{SCSN}_3$, $\text{Br}\cdot\text{SCSN}_3$, and Br_3SCSN_3 . The extreme instability of these substances precludes their exact analysis and description.

VI. THE INTERHALOGENOIDS

The interhalogenoids include cyanazide, CNN_3 (43), cyanogen thiocyanate (6, 8, 13), $\text{CN}\cdot\text{SCN}$, cyanogen selenocyanate (24, 33), $\text{CN}\cdot\text{SeCN}$, and cyanogen azidodithiocarbonate (44), $\text{CN}\cdot\text{SCSN}_3$. Cyanazide is a volatile, crystalline solid melting at 36°C ., which may be prepared by the interaction of sodium azide and cyanogen bromide in suitable solvents:



Cyanogen thiocyanate was first obtained by interaction of cyanogen iodide and silver thiocyanate, and given the name dicyanogen sulfide, $(\text{CN})_2\text{S}$. Its true nature as an interhalogenoid was not recognized until Söderbäck obtained it by the interaction of mercuric cyanide and thiocyanogen:

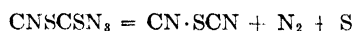


It may also be prepared by interaction of an alkali thiocyanate and cyanogen bromide, although the yields in this case are small (44). It is a beautifully crystalline, extremely volatile solid, melting at 61°C .

Cyanogen selenocyanate, CNSeCN , has been recognized as a

product of decomposition of selenocyanogen. It is a crystalline solid possessing properties very similar to those of cyanogen thiocyanate.

Cyanogen azidodithiocarbonate, $\text{CN} \cdot \text{SCSN}_3$, may be obtained either by interaction of potassium azidodithiocarbonate and cyanogen bromide, or by the interaction of mercuric cyanide and the free halogenoid. It is a white crystalline solid, somewhat volatile, melting with decomposition at 67° . It undergoes spontaneous decomposition in accordance with the equation:



Hydrolysis yields a variety of products, among which are hydrocyanic, cyanic, thiocyanic, sulfurous, thiosulfuric, and sulfuric acids, ammonia, sulfur, and nitrogen. *Ammonolysis* gives principally ammonium azidodithiocarbonate and cyanamide:



The existence of the interhalogenoid, $\text{SCN} \cdot \text{SCSN}_3$, has been postulated to account for the autocatalytic decomposition of azidocarbonyl disulfide (2).

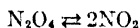
Much other experimental evidence is at hand to support the halogenoid analogy. Compounds such as sulfuryl azide, $\text{SO}_2(\text{N}_3)_2$, (45), carbonyl azide, $\text{CO}(\text{N}_3)_2$ (46), analogous to SO_2Cl_2 and COCl_2 , are known. Thiocyanogen monosulfide and disulfide, $\text{S}(\text{SCN})_2$ and $\text{S}_2(\text{SCN})_2$, have been prepared. Cyanoform, $(\text{CN})_3\text{CH}$, analogous to chloroform is a white crystalline solid melting at 210° . Organic derivatives of the halogenoids have much in common with the corresponding halogen compounds.

Naturally, there are still many loopholes—and the opportunities for further work along this line are innumerable, particularly from the point of view of the organic chemist, be he theoretical or synthetical in aim and purpose. All general halogenic reactions should be applicable to these radicals—and, if found to be otherwise, the cause for their anomalous behavior can unquestionably be found in the structure of the radicals themselves. On the other hand, care must be taken not to apply

halogenic reactions in detail to the halogenoids, as the varied nature of the constituent atoms will be found in many cases to exercise a profound influence on the course of the reactions.

ADDENDUM

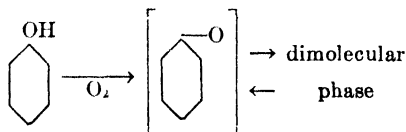
In certain respects *nitrogen tetroxide*, $(\text{NO}_2)_2$, may be thought of as an inorganic radical similar to the halogenoids. Its formation from nitrites, M^+NO_2^- , the existence of compounds like nitroform, $(\text{NO}_2)_3\text{CH}$, nitroxyl chloride, NO_2Cl , etc., seem to bear out such an analogy. The behavior of the tetroxide,



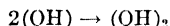
is reminiscent of the behavior of the organic radicals, which dissociate with increasing temperature. Pure nitrogen tetroxide is colorless, but with rising temperature it becomes light yellow, shading gradually into a deep brown color. Decrease in pressure also brings about the same results. These phenomena are accompanied by corresponding changes in molecular weight of the gas so that the existence of $-\text{ON}=\text{O}$ as an unsaturated radical must be assumed. In solution the dissociation of $(\text{NO}_2)_2$ also increases directly with increasing dilution, temperature, etc.

As an inorganic radical, however, it differs markedly from the halogenoids. The latter are colorless in the dimeric form and evidence of their existence in the unsaturated monomeric state is, as yet, lacking. Undoubtedly, these substances do, at the moment of formation, exist as true free (neutral or electrically charged) radicals and it is entirely probable that experimental conditions may yet be ascertained which will give indication of their existence as such to a large extent. It is a curious fact, however, that it is not dissociation, but polymerization which invokes a color change in the case of the halogenoids—exactly the opposite of the behavior of the free *organic* radicals, where appearance of color is supposed to be a sign of dissociation. The fact that the halogenoids do tend to polymerize so readily may be considered as a proof that even the dimeric structure is an unsaturated one. Comparison of inorganic radicals with the so-called “free” organic radicals is not as farfetched as it would

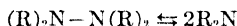
seem at first glance. As in the case of the latter, *tautomerism*⁴ must also be assumed to explain certain chemical reactions. The products formed by oxidation of phenols, the aroxyis (47),



which dissociate into a mono-molecular form, are formally related to the hydroxyl OH radical. Indeed, the latter has been obtained in the dimolecular form as hydrogen peroxide, by electrolysis of certain hydroxides, M^+OH^- (48).



The very fact that the tetra-aryl hydrazines (Wieland) dissociate to yield divalent nitrogen radicals,



would make the momentary existence of the formally analogous NH_2 group as an individual seem within the realm of plausibility. The preparation of hydrazine by interaction of chloramine (formally the interhalogen-halogenoid $\text{Cl}\cdot\text{NH}_2$) and ammonia might easily be construed as a condensation of two such free amine radicals. From an entirely similar point of view, mono-valent sulfur radicals, $-\text{SH}$ or $-\text{SR}$, should be capable of existence.

REFERENCES

- (1) GOMBERG: Chem. Reviews, **1**, 91 (1924)
WALDEN: Chemie der freien Radikalen, Leipzig (1924).
SCHLENK. Sitzungber Berl Akad (1927).
- (2) BROWNE AND CO-WORKERS: J. Am Chem Soc, **45**, 2541 (1923).
- (3) BIRCKENBACH AND KELLERMANN: Ber, **58**, 786, 2377 (1925)
- (4) LIEBIG: Ann. Pharm, **25**, 3 (1838)
- (5) LIEBIG: Pogg Ann, **15**, 541 (1829)
- (6) LINNEMANN: Ann, **120**, 36 (1861).
- (7) SCHNEIDER: Pogg Ann, **129**, 634 (1866)
- (8) SCHNEIDER: J prakt Chem, (2) **32**, 187 (1885)

⁴ See examples given on pages 347, 350, 351 and 352.

- (9) KLASON: J. prakt Chem., (2) **33**, 120 (1886)
- (10) GOLDBERG: J. prakt Chem, (2) **63**, 465 (1901).
- (11) SOMMER: Ber., **48**, 1833 (1915).
- (12) DARROW: Introduction to Contemporary Physics, N. Y., Van Nostrand Co , (1926), p. 20.
- (13) SÖDERBÄCK: Ann , **419**, 217 (1919)
- (14) BJERRUM AND KIRSCHNER C. A , **13**, 1057 (1919); Kgl. Dansk. Videnskab. Selskab., (8), **1**, 76 (1918).
- (15) KAUFMANN AND KÖGLER Ber , **58**, 1553 (1925)
- (16) KERSTEIN AND HOFFMANN Ber , **57**, 491 (1924).
- (17) LECHER AND GOEBEL Ber , **54**, 2223 (1921).
- (18) LECHER AND GRAF: Ber., **59**, 2601 (1926)
- (19) KAUFMANN AND GAERTNER. Ber , **57**, 928 (1924).
- (20) GUTMANN: Z. anal. Chem , **66**, 224 (1925)
- (21) KAUFMANN. Ber. deut. pharm Ges., **33**, 139 (1923).
- (22) KAUFMANN AND KÖGLER Arch. Pharm., **263**, 765 (1925).
- (23) LECHER, WITTEWIT AND SPIER Ber , **56**, 1104 (1923).
- (24) KAUFMANN AND KÖGLER Ber , **59**, 178 (1926).
- (25) LIDOV C. A , **6**, 2368, 2369, 3093, 3094 (1912)
- (26) BROWNE AND HOEL J. Am Chem Soc , **44**, 2315 (1922).
BROWNE AND AUDRIETH. Ibid , **49**, 917 (1927).
SMITH, WILCOXON AND BROWNE Ibid , **45**, 2604 (1923)
- (27) GARDNER Thesis, Cornell University, 1927.
- (28) AUDRIETH AND BROWNE J. Am. Chem. Soc., **49**, 2129 (1927).
- (29) BROWNE AND LUNDELL J. Am Chem. Soc , **31**, 435 (1909).
- (30) WELLS AND PENFIELD Z anorg. allgem Chem , **1**, 442 (1892).
- (31) MATHEWSON AND WELLS Am Chem J , **30**, 431 (1903).
- (32) BIRCKENBACH AND KELLERMANN. Ber , **57**, 2377 (1925)
- (33) VERNEUIL: Ann chim phys., (6) **9**, 326 (1882).
- (34) PRIDEAUX Proc Chem. Soc , **21**, 240 (1905)
- (35) PRIDEAUX. J. Chem. Soc , **89**, 316 (1906).
- (36) STORTENBECKER: Z physik Chem , **3**, 11 (1889).
- (37) TERWOGT. Z. anorg allgem. Chem , **47**, 203 (1905).
- (38) RASCHIG. Ber., **41**, 4194 (1908)
- (39) SPENCER. J Chem Soc , **127**, 216 (1925).
- (40) HANTZSCH Ber , **33**, 522 (1900)
- (41) KAUFMANN AND LIEPE Ber , **57**, 923 (1924).
LECHER AND JOSEPH Ibid., **59**, 2603 (1926).
- (42) BROWNE AND GARDNER: J Am Chem Soc., **49**, 2759 (1927)
- (43) DARZENS Compt rend., **154**, 1232 (1912).
- (44) BROWNE, A. W., AND AUDRIETH, L. F. Unpublished investigations.
- (45) CURTIUS AND SCHMIDT Ber., **55**, 1571 (1922).
- (46) CURTIUS AND HEIDENREICH Ber , **27**, 2684 (1894).
- (47) GOMBERG Chem. Reviews, **1**, 130 (1924)
- (48) MELLOR: Comprehensive Treatise of Inorganic and Theoretical Chemistry, N. Y., Longmans, (1922), **1**, 923.

A SYMPOSIUM ON ATOMIC STRUCTURE AND VALENCE

AN INTRODUCTION

GEORGE L. CLARK¹

Department of Chemistry, University of Illinois, Urbana

The ever-pressing problem of the nature of matter has advanced to the center of chemical and physical interest now more than ever before by the new theories of Heisenberg and Schrödinger, respectively termed quantum mechanics and wave mechanics. The questions on every side are, "What *can* we believe and teach about the atom," and "What next?"

Not so long ago unprejudiced observers might have been wondering if the chemist with his static atom could ever come into agreement or understanding with the physicist's dynamic, infinitesimal solar system or whether indeed there was any effort to meet on a *common* ground by the different means of approach. Both atomic pictures were clearly based on the experiments proving a nuclear structure. Here at least was something in common. Both pictures were useful, the one to explain chemical combination and valence, the other to account primarily for the facts of radiation quantities. On the one hand we have the fruitful contributions of Kossel, Lewis, Langmuir and others; and on the other, the remarkable deductions of Bohr, first for the simple hydrogen atom. While the mechanisms involved in these atomic pictures to account for such chemical facts as stereoisomerism or for the physical facts of emission of radiant energy, seemed remote, some efforts were made to bring chemists and physicists to a common point of view.

Certainly it may be asserted, in spite of discrepancies and dis-

¹ Chairman, Division of Physical and Inorganic Chemistry, American Chemical Society.

agreements, that in the first quarter of the twentieth century, the actual existence of an atomic world became an established fact. Part of the difficulties in details may be ascribed to failure of chemists to test their well-founded conceptions with the facts of physical experimentation, and far too few physicists inquired critically into the facts of chemical combination. So, firmly entrenched each in his own domain, a certain long-range firing of static cubical atoms against infinitesimal solar atoms has ensued, with few casualties and few joint peace conferences.

The position of the Bohr conception has seemed so convincing that perhaps the majority of thinking chemists were coming to accept the dynamic atom, which is fully as capable of visualization as the conceptions placing electrons in definite positions in superposed shells. Valence and stereoisomerism were still difficult but it required only an exercise of faith or a stretch of the imagination to depict *orbits* instead of electrons in space. Thus Main-Smith arrived at spatial representations of electron orbital domains having the general shape, say of a pecan, as the locus of the precession paths of a valence electron moving on the surface of an imaginary solid ellipsoid of revolution with the nucleus at the focus. It is also true that Bohr, Stoner and others made use very largely of the facts of chemistry to extend the dynamic theory to more complex atoms marshalled in the periodic table.

But from the beginning it was extremely difficult to bring the fundamental hypothesis of Bohr into agreement with classical mechanics. Thus wave mechanics was founded in 1924 by Louis de Broglie,—new, though based on such century old conceptions as Fermat's optical principle, Maupertuis' mechanical principle and the familiar Huygens principle, and justifying Bohr's unorthodox mechanics in essence. Schrödinger derived a fundamental differential equation whose constant is an energy term which by its nature can assume only discrete values in the solution. In this way physical problems of atomic mechanics have been reduced to pure mathematics, but the success has been amazing. For the slightly earlier theory of Heisenberg, which had the object of removing from atomic mechanics those quanti-

ties which may not be directly measured by experiment such as the positions and velocities of electrons, now appears to be a necessary consequence of wave mechanics.

Shall now the chemist and the physicist dismiss from their ideas the picture of discrete electrons and quantum orbits and accept only the directly observable quantities characteristic of radiation which must be associated, complexly enough, with two atomic states? Is the atom which is radiating defined by the solution of a differential equation? Shall we say with Dr. Haas that "the second quarter of our century has convinced physicists that a *reality* can be ascribed to the world of atoms only in *another sense* than to that world which reveals itself to man directly through sensuous perception?"

Thus in the face of rapid and remarkable developments the time has seemed ripe to take stock of our conceptions of the nature of the matter. The chemist seeks light on the new wave-mechanics (which somehow has attributes remarkably reminiscent of the old static atom), and on the new experimental facts of radiation and of electrons reflected like waves from crystals. To a singular degree also many physicists are eager to acquaint themselves with knotty chemical problems of valence—for any theory ultimately must be found not wanting in its ability to account for these as well as for spectroscopy.

The Division of Physical and Inorganic Chemistry of the American Chemical Society undertook the task of sponsoring a stock-taking by inviting physicists, and physical, organic and colloidal chemists to contribute to a Symposium on Atomic Structure and Valence at the St. Louis Meeting of the American Chemical Society in April, 1928. So appreciatively was this program received, and so satisfactory seemed the joint stock-taking and mutually educative the coöperation that it has seemed the part of wise service to publish the papers there contributed as a monograph. Each paper represents a different approach to a complex central problem; they all present many debatable subjects, even in the least controversial topic of the nucleus.

Physicists bring the newest contributions of x-rays, of con-

agreements, that in the first quarter of the twentieth century, the actual existence of an atomic world became an established fact. Part of the difficulties in details may be ascribed to failure of chemists to test their well-founded conceptions with the facts of physical experimentation, and far too few physicists inquired critically into the facts of chemical combination. So, firmly entrenched each in his own domain, a certain long-range firing of static cubical atoms against infinitesimal solar atoms has ensued, with few casualties and few joint peace conferences.

The position of the Bohr conception has seemed so convincing that perhaps the majority of thinking chemists were coming to accept the dynamic atom, which is fully as capable of visualization as the conceptions placing electrons in definite positions in superposed shells. Valence and stereoisomerism were still difficult but it required only an exercise of faith or a stretch of the imagination to depict *orbits* instead of electrons in space. Thus Main-Smith arrived at spatial representations of electron orbital domains having the general shape, say of a pecan, as the locus of the precession paths of a valence electron moving on the surface of an imaginary solid ellipsoid of revolution with the nucleus at the focus. It is also true that Bohr, Stoner and others made use very largely of the facts of chemistry to extend the dynamic theory to more complex atoms marshalled in the periodic table.

But from the beginning it was extremely difficult to bring the fundamental hypothesis of Bohr into agreement with classical mechanics. Thus wave mechanics was founded in 1924 by Louis de Broglie,—new, though based on such century old conceptions as Fermat's optical principle, Maupertuis' mechanical principle and the familiar Huygens principle, and justifying Bohr's unorthodox mechanics in essence. Schrödinger derived a fundamental differential equation whose constant is an energy term which by its nature can assume only discrete values in the solution. In this way physical problems of atomic mechanics have been reduced to pure mathematics, but the success has been amazing. For the slightly earlier theory of Heisenberg, which had the object of removing from atomic mechanics those quanti-

agreements, that in the first quarter of the twentieth century, the actual existence of an atomic world became an established fact. Part of the difficulties in details may be ascribed to failure of chemists to test their well-founded conceptions with the facts of physical experimentation, and far too few physicists inquired critically into the facts of chemical combination. So, firmly entrenched each in his own domain, a certain long-range firing of static cubical atoms against infinitesimal solar atoms has ensued, with few casualties and few joint peace conferences.

The position of the Bohr conception has seemed so convincing that perhaps the majority of thinking chemists were coming to accept the dynamic atom, which is fully as capable of visualization as the conceptions placing electrons in definite positions in superposed shells. Valence and stereoisomerism were still difficult but it required only an exercise of faith or a stretch of the imagination to depict *orbits* instead of electrons in space. Thus Main-Smith arrived at spatial representations of electron orbital domains having the general shape, say of a pecan, as the locus of the precession paths of a valence electron moving on the surface of an imaginary solid ellipsoid of revolution with the nucleus at the focus. It is also true that Bohr, Stoner and others made use very largely of the facts of chemistry to extend the dynamic theory to more complex atoms marshalled in the periodic table.

But from the beginning it was extremely difficult to bring the fundamental hypothesis of Bohr into agreement with classical mechanics. Thus wave mechanics was founded in 1924 by Louis de Broglie,—new, though based on such century old conceptions as Fermat's optical principle, Maupertuis' mechanical principle and the familiar Huygens principle, and justifying Bohr's unorthodox mechanics in essence. Schrödinger derived a fundamental differential equation whose constant is an energy term which by its nature can assume only discrete values in the solution. In this way physical problems of atomic mechanics have been reduced to pure mathematics, but the success has been amazing. For the slightly earlier theory of Heisenberg, which had the object of removing from atomic mechanics those quanti-

ties which may not be directly measured by experiment such as the positions and velocities of electrons, now appears to be a necessary consequence of wave mechanics.

Shall now the chemist and the physicist dismiss from their ideas the picture of discrete electrons and quantum orbits and accept only the directly observable quantities characteristic of radiation which must be associated, complexly enough, with two atomic states? Is the atom which is radiating defined by the solution of a differential equation? Shall we say with Dr. Haas that "the second quarter of our century has convinced physicists that a *reality* can be ascribed to the world of atoms only in *another sense* than to that world which reveals itself to man directly through sensuous perception?"

Thus in the face of rapid and remarkable developments the time has seemed ripe to take stock of our conceptions of the nature of the matter. The chemist seeks light on the new wave-mechanics (which somehow has attributes remarkably reminiscent of the old static atom), and on the new experimental facts of radiation and of electrons reflected like waves from crystals. To a singular degree also many physicists are eager to acquaint themselves with knotty chemical problems of valence—for any theory ultimately must be found not wanting in its ability to account for these as well as for spectroscopy.

The Division of Physical and Inorganic Chemistry of the American Chemical Society undertook the task of sponsoring a stock-taking by inviting physicists, and physical, organic and colloidal chemists to contribute to a Symposium on Atomic Structure and Valence at the St. Louis Meeting of the American Chemical Society in April, 1928. So appreciatively was this program received, and so satisfactory seemed the joint stock-taking and mutually educative the coöperation that it has seemed the part of wise service to publish the papers there contributed as a monograph. Each paper represents a different approach to a complex central problem; they all present many debatable subjects, even in the least controversial topic of the nucleus.

Physicists bring the newest contributions of x-rays, of con-

temporary ultra-spectroscopy, of wave-mechanics. Chemists discuss the nucleus in its every aspect, make a genuine liason attempt to account for valence, outline with a challenge the old and new problems and theories of organic compounds, and present new physical methods of measuring chemical bonds, and strange phenomena in supermolecular states of matter. The Symposium is a cross-section of contemporary thought and achievement—heterogeneous, tentative, often discordant, at once clear and abstruse—and yet remarkably inspiring in the genuine effort to organize widely different fields of information into a true *general* science of matter and to press on to one of the great goals of intellectual endeavor. It is sincerely hoped that these papers may educate, refresh and stimulate in the general study and the seminar; and that they may be the documentary evidence of the welding in 1928 of Chemistry and Physics on a great common ground.

RECENT PROGRESS IN THE STUDY OF THE STRUCTURE OF THE NUCLEUS

S. C. LIND

School of Chemistry, University of Minnesota

Sir Ernest Rutherford first demonstrated by means of the bombardment of atoms by α particles that nearly all of the atomic mass is concentrated in the nucleus. The work of Lewis, Langmuir, Kossel, Bohr, and many others has demonstrated that *chemical* union is effected by means of extra-nuclear electrons in outer or valence orbits. But it should not for this reason be concluded that the chemist has no concern with the nucleus, that on account of its minuteness or inaccessibility that it has a structure of interest only to physicists and mathematicians. It is the nucleus which determines the atomic number and hence the ultimate identity of each atom, it determines the existence of isotopes, and all of the radioactive properties. Finally whatever chances there be of transmuting or of disrupting the atom must depend on our ability to reach the nucleus and to bring about changes in its structure.

Owing to the *infinitesimal* size of the nucleus which even for the heaviest atoms has a radius of the order $3 - 6 \times 10^{-12}$ cm. or a volume of the order 10^{-35} cm.³ and to the large number of structural units that must be contained in it, which for the uranium nucleus is 238 protons and 146 nuclear electrons, it was at first difficult to conceive of its having anything but a closely packed structure. One of the early popular conceptions was that of bricks representing the protons and helium nuclei held together by electrons interspaced like mortar. Indeed a structure suggested by Rutherford (1) for the close packing of helium nuclei in layers with one electron at the center of each cube formed by 8 adjacent helium nuclei gave the correct atomic number for various elements from titanium (22) to platinum (78).

We have, however, numerous indications that a closely packed static structure is not capable of interpreting even the data that we already possess regarding the properties of the nucleus. The fact that rays and high velocity particles are emitted from the nucleus indicates that it is partly at least a dynamic, not a static, system. Just as the multiplicity of the spectroscopic lines had predicted long before we had any definite ideas about electronic structure of the atom that it must have a fairly complicated structure which we now know as the orbital electronic structure, so do the α , β , and γ radiations from the nucleus predict a complex structure. And just as the application of quantum principles to orbital electrons has elucidated the structure and the mechanism of their behavior, so also it appears that quantum principles will be the surest guide in solving the structure of the nucleus, and in interpreting its radioactive instability.

Just as in spectroscopy, the multiplicity of lines was early an embarrassment but later furnished the essential proof of the quantum theory of atomic structure, so we have an accumulation of physical and chemical data awaiting interpretation through a complete theory of the nucleus. On the one hand we have the occurrence or lack of isotopes for each atomic species. Moreover, the occurrence of isobares proves that fundamental differences exist which can only be attributed to structure. The work of Aston has furnished an abundance of data, the precision of which he has recently raised to a degree capable of standing very severe interpretative tests.

Among the ordinary atoms we have then a structure to be interpreted by the distribution of protons and electrons in the nucleus to make up any given atom or isotopes of it. In the radioactive atoms we have also a wealth of experimental evidence. First we know that it is only the heaviest atoms, with the extreme exceptions of potassium and rubidium, which possess radioactive properties. We know also with considerable accuracy the life periods of the various radioactive elements, the type and energy of the radiations emitted by each one and the relation of life period to energy, such as the Geiger-Nuttall relation for range of α particles to the life period of these emitters. Further-

more, we know that the only atomic particles spontaneously emitted from the nucleus of the radioactive atoms are α particles (helium nuclei) and that the only particles which have hitherto been ejected from the light atoms by α ray bombardment are protons. (But since there is such a wide difference in mass between the radioactive elements and the light ones which can be disrupted by α particles, no present conclusion about the general predominance of protons or helium nuclei in certain regions of all nuclei would be justified.)

The idea of quantized states in the nucleus arises from the marked equality of energy possessed by α particles of the same variety, that is by those from the same element. The α particles of polonium have been shown by Irene Curie to have remarkably uniform initial velocity while the same has been demonstrated by Briggs for the much swifter α particles from Ra C'. If the particles were in each case before their emission moving in or about the nucleus with equally uniform velocity, then evidently we have indications of definite and probably of quantized orbits. γ rays from a given atom also exhibit a high degree of homogeneity of wave length. But β particles, on the other hand, are emitted from the same element with velocities varying continuously through a wide range.

It was previously thought that this variation in the speed of β particles could be explained by assuming their emission at speeds corresponding to voltages all the way from 40,000 to 1,050,000 volts to be due to partial expenditure of the initial energy of the swifter β rays before they emerge from the atom. Ellis and Wooster (2) have, however, recently measured the heat evolved in the disintegration of RaE and found it equal to about 350,000 electron-volts, which corresponds to the *average* energy of β ray emission not to the maximum 1,000,000 volts. Therefore they conclude by analogy that the initial β rays from all atoms are truly heterogeneous and that they originate in a part of the atom where they are in random rather than in quantized motion.

Rutherford (3) has recently brought the knowledge of α ray emission to bear on a theory of nuclear structure. His result is

that the α particle before emission is a neutral particle revolving in quantized orbits about the nucleus. Its electrons are not in the normal helium orbit but held very close to the nucleus of the satellite. The force holding the neutral atom in its orbit is the attraction due to the distortion or polarization by the central nuclear field. When the α particle is expelled it leaves its two electrons behind, carrying away a double charge as is known from the atomic number of the resulting atom which is lower by two units than that of the parent atom. The two electrons left behind move in closer to the central nucleus from which they are probably emitted later as β particles. Frl. Meitner earlier suggested the neutral α satellite on account of the frequent succession of one α ray by two β ray emissions.

Rutherford (2) developed an equation for the quantized emission of α particles and assigned by trial a quantum number (using half quantum numbers also) to each radioactive element. He evaluated the three universal constants involved in the equation from a standard atomic number 84. The calculated energies of the different sets of α particles agree with observation within a fraction of 1 per cent. The quantum numbers fall in the range 14 to 30, corresponding to distances from the center between 1.9×10^{-12} and 4.5×10^{-12} cm. Probably satellites can exist for atomic numbers as low as 15 but become frequent for atomic numbers about 30, and since we can associate isotopes with the possibility of additional neutral satellites which do not alter the atomic numbers, it is interesting that isotopes become numerous for atomic numbers about 29. The α ray evidence applies only to satellites of mass 4. Unfortunately as Rutherford points out there are no similar data from which estimation may be made in regard to the possibility of adding satellites of mass 2, which are known to play as important rôle in the structure of isotopes and which Harkins has found to be involved in the structure of all atoms.

From the foregoing considerations we can draw the following present picture of the nucleus: a very small compact central portion with a net positive charge, surrounded by a region of electrons in random motion, and finally outer neutral helium

satellites revolving in quantized orbits from which they may be ejected from the heavier atoms as α particles of definite energy and speed.

Although the scheme of Rutherford is based wholly on electrostatic forces, it has also been suggested that the nucleus is held together by magnetic forces or a combination of electric and magnetic forces. D. Enskog (4) last year put forth a theory taking into account both forces, by means of which he calculates very accurately not only the energies of the α particles but of the β particles as well.

Many attempts have been made to disrupt or change the nucleus. The only success has been had with α particles, which have been shown to change the nucleus by knocking out a proton and by adding the helium nucleus to the nucleus of the bombarded atom. These results will doubtless contribute to the ultimate solution of atomic structure.

Attempts using less energetic agents have not been successful. Soddy first pointed out that to lower the atomic number of an atom by *one*, it would only be necessary to shoot an electron into the nucleus thus neutralizing one unit of positive charge. Various claims of transmutation have not been substantiated. It is evident that the means that have been employed to change the nucleus either by means of high speed electrons or of high temperature have not been sufficiently energetic. The goal appears by no means an impossible one, and from the reports of high voltage electrons soon to be employed there is a prospect of its attainment in the not distant future.

REFERENCES

- (1) RUTHERFORD, J. Franklin Inst., 198, 725 (1924).
- (2) ELLIS AND WOOSTER, Proc Roy. Soc (London), 117A, 109-123 (1927).
- (3) RUTHERFORD, Phil. Mag. (7), 4, 580-605 (1927).
- (4) ENSKOG, Z. Physik., 45, 852-868 (1927)

THE SYNTHESIS OF ATOMS, THE WHOLE NUMBER RULE, AND THE PERIODIC SYSTEM OF THE ATOMIC SPECIES¹

WILLIAM D. HARKINS

Department of Chemistry, the University of Chicago, Illinois

I. THE DISINTEGRATIVE SYNTHESIS OF ATOMS

1. Introduction

In the decade between 1860 and 1870 a study of the relations between the chemical and physical properties of the elements and their atomic weights led to the formulation of what is commonly designated as the periodic system of the elements. The development of this system was due to de Chancourtois, Newlands, Mendeléeff, Meyer, and a number of others. Recent investigation concerning what is commonly called the structure of the atom seem to indicate that the part of the atom to which this system is particularly related is the non-nuclear portion. It must be admitted that while in recent years much has been learned of the properties of materials and of the behavior of very small amounts of substances, particularly dilute gases, almost nothing is known of the *structure* of any atom. The idea concerning this structure which seems best supported by the experimental evidence is that the portion of the atom which contains nearly all of its mass (the nucleus) is extremely small in comparison with the entire atom, and is charged with positive electricity. Nothing is known as to the structure of either this nucleus or the more diffuse system (presumably charged with negative electricity) which seems to surround it, except that there seems to be a cer-

¹ Since the preceding paper by Dr S. C. Lind gives a summary of the work on the nucleus with the exception of that carried out in Kent Chemical Laboratory, the present paper is restricted to the work of this laboratory and that upon which it is based.

tain, unknown, amount of discontinuity involved. That this paucity of knowledge of the structure of the non-nuclear atom is not due to lack of attention to the subject by physicists and chemists is made evident by the literature of the last decade.

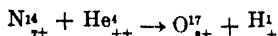
The nuclear portion of the atom seems to be much more shielded from observation, so the number of known facts as to its characteristics is relatively extremely small. It is therefore of importance to determine the general relations or laws which may be deduced from those facts which are available. In 1914 the writer made a careful study of the atomic weights of the elements and of the radioactive atomic species, and of their abundance in meteorites and on earth. A number of fundamental laws, which seemed to be related to the composition of the nucleus, were found to emerge, and these laws or relations appeared to indicate that certain steps may be involved in the building of atoms. Thus it seemed almost certain that atoms may grow heavier by the addition of helium atoms, or speaking more technically, that the nuclei of atoms may grow heavier by the addition of α -particles.

The work described in the section which follows was begun in 1921 in order to obtain evidence in favor of this idea. The relations or laws which gave rise to this work are outlined in the later sections. They constitute a periodic system of the atomic species which seems to be related to the characteristics of atom nuclei.

2. The application of Wilson ray tracks to the study of the disintegration of the atom

About seven years ago Harkins and Ryan (4) developed and applied a procedure capable of revealing new details concerning the processes which occur in the disintegration of the light atoms. This procedure, adopted by Blackett (5) and by Harkins and Shadduck (6), has shown that what was supposed by Rutherford to be the disintegration of the atom, is actually a combined synthesis and disintegration, which may be called a disintegrative-synthesis. In such a process a fast doubly charged helium atom (α -particle) collides and unites with the nucleus of a nitrogen

atom of mass 14 and charge 7+. It may be assumed that this forms a nucleus of a fluorine atom of mass 18 and charge 9+, which explodes almost immediately and emits a long range proton (p or H⁺). It may even be possible that the hydrogen particle (H⁺) is emitted while the alpha particle is entering the nucleus, so that the life of the fluorine nucleus may be almost zero. The process may be considered as a metathesis or double decomposition of atom nuclei.



The initial discovery of the disintegration of light atoms was made by Rutherford, who designed a method in which the scintillation apparatus of Crookes was so modified as to show when long range H-particles or protons are ejected from atom nuclei bombarded by alpha particles.

Harkins and Ryan (4) considered that similar modification of the cloud track apparatus of Wilson would make it capable of exhibiting a greater part of the mechanism of disintegration.

That the photography of Wilson cloud tracks had not been used earlier for the purpose is undoubtedly due to the fact that the disintegration of atoms under bombardment is such a rare event that this method had not appeared to others as suitable. It will be shown later, too, that the change in the source of alpha rays, from the polonium applied earlier to a source such as ThC or RaC, which emits fast rays, was essential to the success of the method.

By the use of the extraordinarily fast rays from thorium C', Harkins and Ryan obtained 11,000 photographs in nitrogen and 20,000 in argon. That no H-disintegrations of the ordinary type were obtained in this work was partly due to the fact that a considerable part of the work was carried out in argon, which disintegrates less often than nitrogen, and partly to a statistical accident.

3. The disintegrative-synthesis of atoms

That the disintegration of nitrogen is accompanied by an atomic synthesis is shown by the photographs of both Blackett (5), and

Harkins and Shadduck (6), but the success of the work in both cases was due to the use of the method of Harkins and Ryan without any essential or important change.

The work reported here is a continuation by Harkins and Shadduck of the work of Harkins and Ryan by the use of the same, but modified, apparatus, which consisted in a modified Shimizu-Wilson cloud track apparatus. The number of photographs taken by us is 34,000 with an average of slightly more than 12 α -tracks each, so that 270,000 tracks of 8.6 cm., and 145,000 tracks of 4.9 cm. range were obtained.

Figure 1A shows the arrangement in a plane of the tracks in one of the cases in which a fast α -particle attaches itself to the

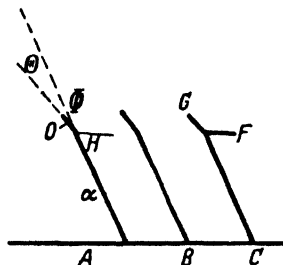


FIG. 1

A, a disintegrative-synthesis of oxygen (*O*) from nitrogen and helium. *B*, the same disintegrative-synthesis with the track of the proton (*H*) omitted. *C*, an ordinary collision between the alpha-particle and the nucleus of an atom of nitrogen.

nucleus of a nitrogen atom. This, as has been stated, forms momentarily the nucleus of an atom of fluorine which disintegrates almost at once to give a fast H-particle or proton, and the nucleus of a heavier atom, presumably that of oxygen of mass 17. All three tracks lie in a plane. The angle between the direction of the track of the α -particle and that of the H-particle is 118° . The photograph indicates that the track of the proton has only about *one-sixth* to *one-tenth* the intensity of that of the α -particle or of the oxygen atom. Figure 1B illustrates what would be formed provided the track of the H-particle should be so faint as to be invisible. Figure 1C illustrates the marked differ-

ence in the appearance when the α -particle rebounds in the case of an ordinary (but extremely rare) collision. Here the thin track is not present and the angles are noticeably different. Figure 2 is a reproduction of a photograph of the disintegrative-synthesis or atomic metathesis represented by figure 1A.



Fig. 2 TWO VIEWS AT RIGHT ANGLES OF THE NON-ELASTIC COLLISION REPRESENTED IN FIGURE 1, A

At the left of each view an α -particle impinges upon the nucleus of a nitrogen atom and unites with it. An extremely faint straight track extends horizontally to the right of the fork and is due to a positron or H-particle emitted at high speed. This has a range of 19.6 cm. in air at 15" and 760 mm. pressure. The track of the oxygen atom which is formed slants upward to the left from the fork. The left-hand view is overexposed as a result of an irregularity in the lighting, but the track of the H-particle is easily visible on the original negative.

Table 1 gives the characteristics of this fork.

The only evidence that the inelastic collision represented by fork 2 is a disintegrative synthesis, is the appearance of the fork. The thin track has exactly the appearance given by hydrogen.

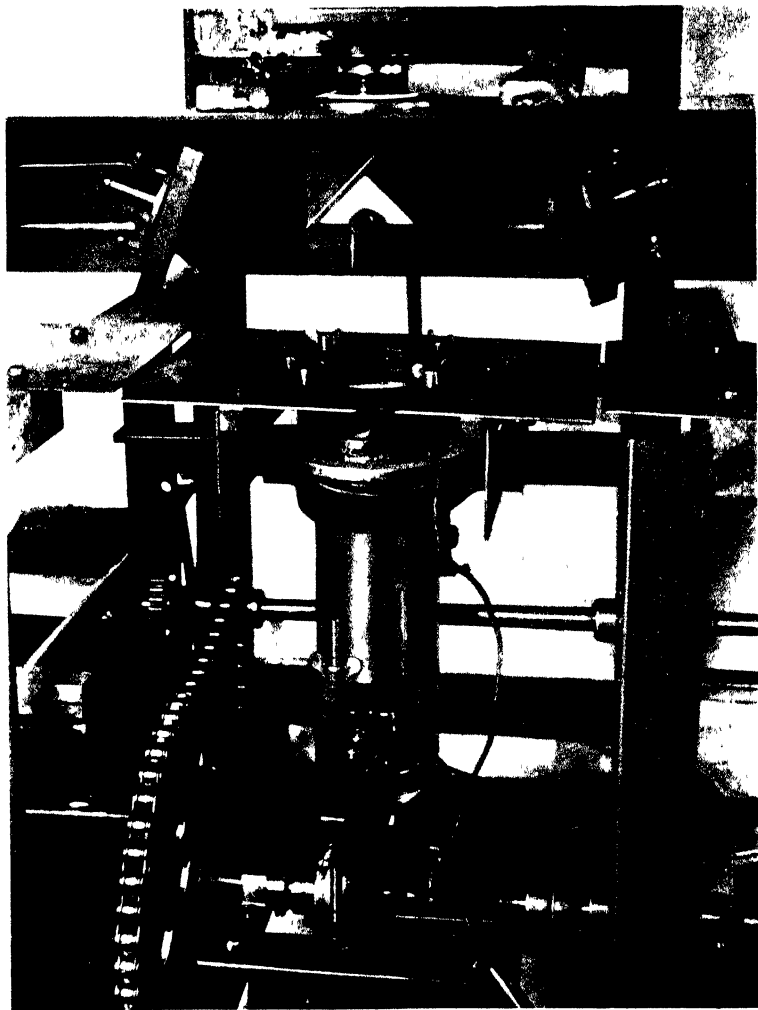


FIG. 3 MODIFIED WILSON (SHIMIZU) APPARATUS FOR THE PHOTOGRAPHY OF ALPHA-RAY TRACKS

However, measurements on the fork cannot be exact on account of curvature near the point of impact

Figure 3 represents the ionization chamber used in the work,

and figure 4 the natural disintegration of thoron to give thorium A, and the subsequent disintegration of thorium A to give thorium B (Lead). As in all other cases, the two views are taken at right angles to each other.

4. The mesh plate method for the measurement of alpha- and beta-ray tracks

The measurements of the tracks were made by a new method which may be called the "Mesh Plate Method." It consists in

TABLE I
Characteristics of forks, which represent the disintegrative-syntheses of atoms

	ϕ	θ	R_α	$V_\alpha \times 10^{-8}$	$V_p \times 10^{-8}$	R_p
Fork 1	118° 32'	15° 21'	6.4	1.86	2.7	19.6
Fork 2	24° 30'	6.3	4.1	1.57		
				$V_n \times 10^{-8}$	R_n	$E_0 - E_1$
Fork 1				0.534	0.43	0.89

R_α = remaining range of α -particle, the distance the α -particle would have traveled beyond the point of the fork if the collision were not incurred

V_α = velocity of the α -particle in centimeters per second immediately before

V_p = velocity of the proton immediately after its escape from the heavy nucleus

R_p = range of the proton, or H-particle

V_n = velocity of the oxygen nucleus immediately after the escape of the proton

R_n = range of the nucleus of the oxygen atom

$E_0 - E_1$ = kinetic energy of the H-particle plus that of the oxygen atom, divided by the kinetic energy of the α -particle immediately before impact. $1 - E_0 - E_1$ represents the loss of kinetic energy due to the non-elastic impact, and is presumably stored up in the oxygen nucleus of mass 17

the establishment in the ionization chamber, which consists of a cylinder of optical glass, of a system of coordinates in space. A circular plate of yellow brass was made of such a size as to fit the cylinder with a clearance of about 0.5 mm. This was ruled on the milling machine into squares of 1 sq. mm. area by lines 0.1 mm. in diameter. These lines were then filled with tin solder and the brass surface blackened by the use of ammoniacal am-



FIG 4 TWO VIEWS OF THE NATURAL DISINTEGRATION OF THORON TO GIVE THORIUM A, AND THE SUBSEQUENT DISINTEGRATION OF THORIUM A TO GIVE THORIUM B
(Views at right angles)

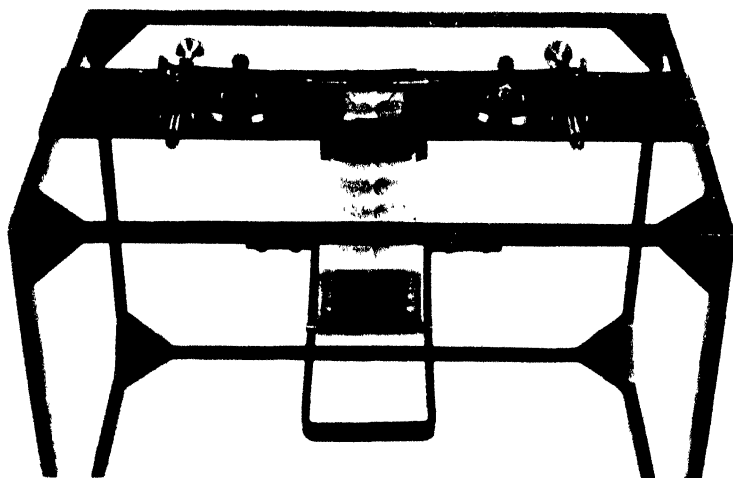


FIG 5 APPARATUS FOR THE MEASUREMENT OF ATOM RAY-TRACKS

monium carbonate. This gave very uniform, silver-white lines against black on an entirely plane surface, since the surface was planed prior to the blackening.

The mesh plate is then set on the brass piston, which closes the bottom of the ionization chamber and fits inside a cylinder of brass below the glass. The mesh plate is set in an exact position by the use of a stand with three legs fitted with screws. The bottom of each leg fits a fine screw hole in the top of the piston. By the use of the fixed mirrors, used in obtaining two photographs at right angles of each set of α -tracks, two photographs at right angles, are taken of the mesh plate. The piston is then raised a distance of 0.02 cm. and another pair of photographs taken. This operation is repeated until the top of the chamber is reached.

In order to find the position in space represented by any point in the photographs, it is only necessary to find the point in both photographs. The film which shows the point is then placed in a special apparatus (fig. 5) which fits the film with considerable exactness. This is provided with two movable and very sharp pointers. One of these sharp pointers is put on each projection of the point on the film, and fastened in place, except that a vertical movement is possible. The film is then taken out, and the film with the mesh plate images is inserted. This is moved until the reading given for the position of each pointer is the same on one projection as on the other. The reading gives the x and y coördinates, and the number of the film, the z -coördinate, of the point. This method gives the position of a point in space to within 0.1 mm. The use of this method of measurement eliminates errors due to the lens, which was, however, a Taylor-Cooke-Hobson lens of the best grade. Also errors due to the glass at the top of the cylinder, and to the glass of the mirrors are eliminated. Nevertheless all of these were made optically plane in Ryerson Physical Laboratory. If the lines are too short or too highly curved, other methods are used. The mesh-plate method has the great advantage that it allows not only for all errors due to the use of the lens but also for differences in magnification due to the use of two views at right angles.

5. Minimum velocity of the α -particle in a disintegrative-synthesis

The minimum velocity for the α -particle which has thus far been shown definitely to give rise to an atomic disintegrative-synthesis corresponds to a remaining range of 6.2 cm. This corresponds to the energy given by 7×10^6 electron-volts or to the velocity given to the α -particle by 3.5×10^8 volts. However, Fork 2 seems to represent a synthesis in which a much lower velocity (1.61×10^9 cm. per second), which corresponds to 5.4×10^6 electron-volts, was effective. This is only slightly higher than the initial velocity of the α -particles from polonium (1.587×10^9 cm. per second, or 5.224×10^6 electron-volts), so it is probable that it is possible to obtain such a disintegrative-synthesis even when polonium is used as the source of α -rays, but polonium is altogether unsuitable in the photographic method as applied in this laboratory, that is if a photograph of all of the rays of the fork are obtained. The method of Holubek (7) is of an entirely different order, since it gives no indication that a synthesis has occurred, and is more nearly analogous to the ordinary scintillation method. Up to the present, the photographic method has not demonstrated the occurrence of disintegrations which give short H-rays.

II. THE WHOLE NUMBER RULE OF HARKINS AND WILSON (1)

In 1815 Prout, on the basis of extremely inexact data, announced the relation that if the atomic weight of hydrogen is taken as unity, the atomic weights of the other elements are exactly whole numbers.

Table 2 gives the atomic weights according to Prout, and also as given in the latest tables. It is apparent that deviations from the true values as high as 5.9 units for carbon, 15.8 units for sulphur, 19.7 units for calcium, 27.4 units for iron, and 66.3 units for barium, together with the general smaller deviations, indicate that there was no real basis at the time for the whole number rule with hydrogen taken as unity.

While Prout's Whole Number Rule is thus shown to be altogether without justification, the idea which he based upon these

faulty data, that hydrogen is the fundamental element, or protyle, is in accord with the views of the present time. This is a good illustration of the fact that a good theory may be based upon altogether unsound data.

The *modern* whole number rule is more properly called the law of the approximate constancy of the packing effect.

According to this rule *the atomic weight of any atomic species (other than hydrogen) is very close to a whole number on the basis of 16 as the atomic weight of oxygen.*

TABLE 2

Prout's table of the more accurately determined (1815) atomic weights
Atomic weight (H = 1)

ELEMENT	PROUT (1815)	VALUE 1928	DIFFERENCE
H	1	1 -	0 (Standard)
C	6	11 907	5 907
N	14	13 900	-0 1
P	14	30.787	16 787
O	16	15.876	-0 124
S	16	31 819	15.819
Ca	20	39 70	19 70
Na	24	22 819	-1.181
Fe	28	55.41	27 41
Zn	32	64.87	32.87
Cl	36	35 183	-0 817
K	40	38 793	-1 207
Ba	70	136.31	66.31
I	124	125.950	1 950

Since the atomic weight of hydrogen is, as nearly as is known, equal to 1.0078, this indicates that the packing effect in a synthesis of any other element from hydrogen is very nearly equal to a loss of mass of 0.77 per cent. When this law was first published it was considered that the packing effect would be constant to within about 0.1 per cent.²

² It was shown later (fig. 22) that the atoms of odd atomic number are less stable than those of even atomic number, so it is not surprising that for the lightest atoms of odd atomic number the packing is 0.57 and 0.60 per cent (lithium

A few quotations from the first paper concerning the whole number rule, are given below.

Not only is the variation of the atomic weight from a whole number a negative number, but in addition its numerical value is nearly constant, the average value for the 21 elements being 0.77 per cent, while the six elements from boron to sodium show values of 0.77, 0.77, 0.70, 0.77, 0.77, and 0.77 per cent. The deviation is therefore not a periodic, but a constant one.

It has formerly seemed difficult to explain why the atomic weights referred to that of oxygen as 16 are so much closer to whole numbers than those referred to that of hydrogen as one, but, the explanation is a very simple one when the facts of the case are considered. The closeness of the atomic weights on the oxygen basis to whole numbers, is indeed extremely remarkable. If an oxygen atom is a structure built up from 16 hydrogen atoms, then the weight according to the law of summation should be 16 times 1.0078 or 16.125. The difference between 16.125 and 16.00 is the value of the packing effect, and if this effect were the same for all of the elements, except hydrogen, then the choice of a whole number as the atomic weight of any one of them, would, of necessity, cause all of the other atomic weights to be whole numbers. Though this is not quite true, it is seen that the packing effect for oxygen is 0.77 per cent, which is the average of the packing effects for the other 21 elements considered. Therefore, those elements which have packing effects equal to that of oxygen will have whole numbers for their atomic weights, and *since the other elements show nearly the same percentage effect, their atomic weights must also lie close to whole numbers.*

It has usually been assumed, and without any really logical basis for the assumption, that if a complex atom is made up by the union of simple atoms, the mass of the complex atom must be exactly equal to the masses of the simple atoms entering into its structure. Rutherford, from data on the scattering of α -rays in passing through gold leaf, has calculated an *upper* limit for the radius of the nucleus of a gold atom as

of atomic weights 6 and 7) as found by Costa. In these two special cases the variation from the whole number relation is abnormally large: 0.2, and 0.17 per cent.

Also it would not be surprising if the packing is found to be somewhat smaller for uranium and thorium and adjacent elements, than for elements of half their atomic weight.

3.4×10^{-12} cm. The mass of this relatively heavy atom is, according to this calculation, practically all concentrated in this extremely small space, which is so small that it could no longer be expected that the mass of such a nucleus, if complex, would be equal to the sum of the masses of its component parts. In fact, since the electromagnetic fields of the electrons would be so extremely closely intermingled in the nucleus, it would seem more reasonable to suppose that the mass of the whole would not be equal to the sum of the masses of its parts. The deviation from the law of summation cannot be calculated on a theoretical basis, but it can easily be determined from the atomic weights, if the assumption is made that the heavier atoms are condensation products of the lightest of the ordinary elements, that is of hydrogen. This deviation expressed in terms of the percentage change, is what has already been determined, and designated as the packing effect.

The closeness to which a positive and a negative electron would have to approach to give a decrease of mass equal to 0.77 per cent, or the average value of the packing effect, is found by calculation to be to a distance of 400 times the radius of the positive electron. This case does not correspond to any element actually known, for the simplest of the atoms considered, helium, may be supposed to have a nucleus built up from four hydrogen nuclei and two negative electrons. However, the magnitude of the effect seems to be of the order which would be expected.

1. The atomic weights of the first 27 elements, beginning with helium, are not multiples of the atomic weight of hydrogen by a whole number, as they would be if Prout's original hypothesis in its numerical form were true. This may be expressed by the statement that the atomic weights on the hydrogen basis are not whole numbers. However, when these atomic weights are examined critically it is found that they differ from the corresponding whole numbers by a nearly constant percentage difference, and that the deviation is negative in sign, with an average value of -0.77 per cent.

2. This percentage difference has been called the packing effect, and it represents the decrease of weight, and presumably the decrease of mass, which must take place if the other atoms are complexes built up from hydrogen atoms. The regularity in this effect is very striking, the values for a number of the lighter atoms being as follows: He, -0.77 ; B, -0.77 ; C, -0.77 ; N, -0.70 ; O, -0.77 ; F, -0.77 ; and Na, -0.77 per cent, while the average value for the first 27 elements is -0.77 per cent.

As has been seen, there are 27 atomic weights distributed over 59 units of atomic weight. The greatest common divisor of the whole numbers corresponding to the atomic weights is one. *The atomic weights are therefore such that numerically they seem to be built up from a unit of a mass of one, and the probability results seem to show that this unit of mass must be very close to 1.000, expressed to three decimal places.*³

Now that certain elements have been found to exist in isotopic forms, it becomes apparent that still other elements may do the same in cases which have not been recognized, so that in dealing with any single species of element it is uncertain whether this is an individual with respect to its atomic weight. The great regularity with which the elements follow the relationships given in these papers, up to an atomic weight of 59, suggests that with the exception of the cases of neon, silicon, magnesium, and chlorine, isotopes probably do not exist to any large extent for any of these elements, if they exist at all.

III. THEORETICAL: THE BUILDING OF OTHER ATOMS FROM HYDROGEN

1. Introduction

The fact that the atomic weights of the atomic species are whole numbers (to within 0.1 per cent in general) suggests that all of the more complex atoms are built up from hydrogen, but that in the formation of the complex structure the atomic weight of the hydrogen is changed from 1.0078 to 1.000 (± 0.001). Since, however, hydrogen is supposed to consist of a positive electron (proton) and a negative electron, this is equivalent to the statement that complex atoms are built up from protons and electrons.

In such complex atoms the masses of the protons are not independent, but the average mass of the proton-electron pair in any atom may be considered as 1.000 ± 0.001 ($1.649 \times 10^{-24}\text{g.}$), if the atomic weight of oxygen is taken as 16.

Since only about 1/1840 of the mass of the hydrogen atom is attributed to the electron, the proton is supposed to contain nearly all the mass; that is, the mass is found to be associated almost wholly with the positively charged particles.

³ Not italicized in the original.

As nearly as is known the absolute magnitude of the positive charge of the proton is equal to that of the negative electron. Any complete atom is apparently electrically neutral, and an atom which is not electrically neutral is designated as an ion.

2. General stability relations of atom nuclei

The existence of atoms and of atomic species is highly dependent upon the numerical values of certain experimental

TABLE 3

Experimental and theoretical significance of fundamental variables related to atomic stability

VARIABLE	EXPERIMENTAL SIGNIFICANCE	THEORETICAL SIGNIFICANCE
P	Atomic weight of atomic species or protonic number	Number of protons in the nucleus (respectively in the atom)
Z	Atomic number	Net positive charge on the nucleus
N	$P - Z$ Electronic number	Number of negative electrons in the nucleus
n	$P - 2Z$ $2N - P$ $N - M$ Isotopic number	
$\frac{N}{P}$	$\frac{P - Z}{P}$ Ratio of electronic to protonic number	Ratio of negative to positive electrons in the nucleus

quantities as listed in table 3. Since existence depends upon stability, it seems apparent that the stability of atoms is dependent upon the magnitudes of the same quantities.

It is important to emphasize that N is an entirely experimental number, since it is the difference between the atomic weight and the atomic number, both of which are determined by experiment. The same is true of the isotopic number (n). It may be noted

that from the experimental standpoint all of the four variables P , Z , N , and n , are very close to whole numbers, while from the theoretical standpoint they are exactly whole numbers.

The *protonic number* may now be defined as the *whole number* which is nearest to the mass number (or atomic weight) of the atomic species. The *electronic number* is the *whole number* which is nearest to the difference between the protonic number and the atomic number ($N = P - Z$).

3. *The nucleus as a proton-electron aggregate*

If the nucleus is considered as built up from protons and electrons, without any reference to secondary groups or structure, some interesting relations emerge.

All of the protons of the atom, and usually half the electrons, go into the nucleus, while the other half of the electrons make up a more diffuse system, which is responsible for the emission and absorption spectra in the visible, ultra-violet, and x-ray regions of the spectrum. In the case of only about 18 per cent of known atoms, *more than half* of the electrons enter the nucleus, so for these atoms the ratio N/P is greater than $1/2$. The numerical relations involved are discussed in Section IV.

IV. GENERAL NON-PERIODIC RELATIONS CONCERNING ATOM NUCLEI

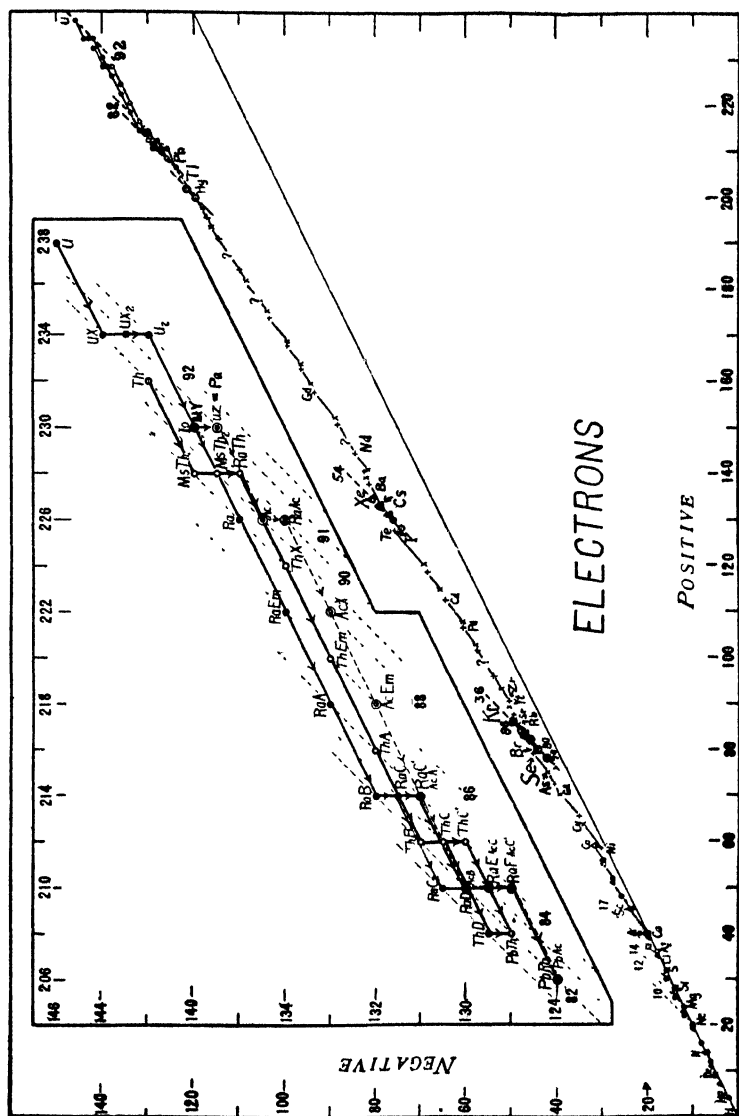
1. *Stability and the ratio of negative to positive nuclear electrons* (3)

One of the most fundamental of all of the relations or laws which concern the stability of atom nuclei is as follows:⁴

No species of atoms (other than hydrogen) which has been found to exist either on earth or in the meteorites has a value for the ratio of negative to positive electrons $\left(\frac{N}{P}\right)$ of less than $1/2$ or 0.5 .

It is remarkable, too, that in most of the atoms which exist the

⁴ This relation was used by the writer as a basis for the development of the "Hydrogen-Helium Theory," J. Am. Chem. Soc., **37**, 1383 (1915), **39**, 856 (1917), and is discussed in considerable detail in *ibid*, **42**, 1964-75 (1920), Phil. Mag., **42**, 308 (1921).



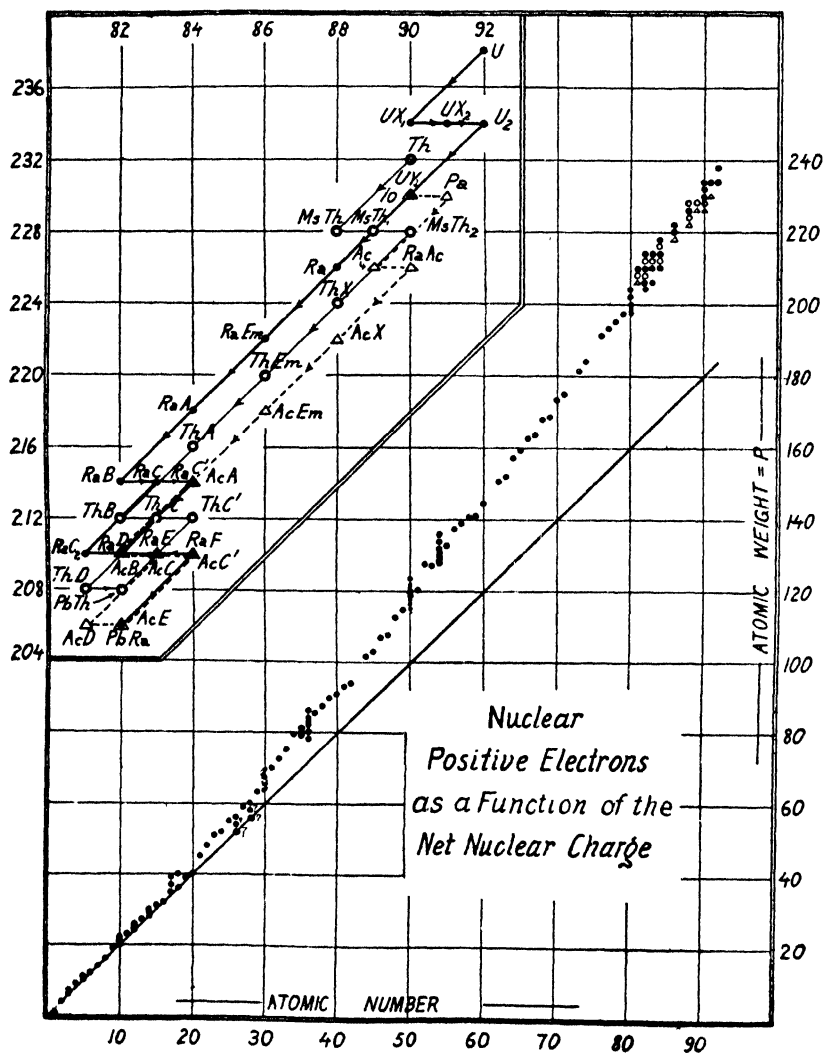


FIG. 7

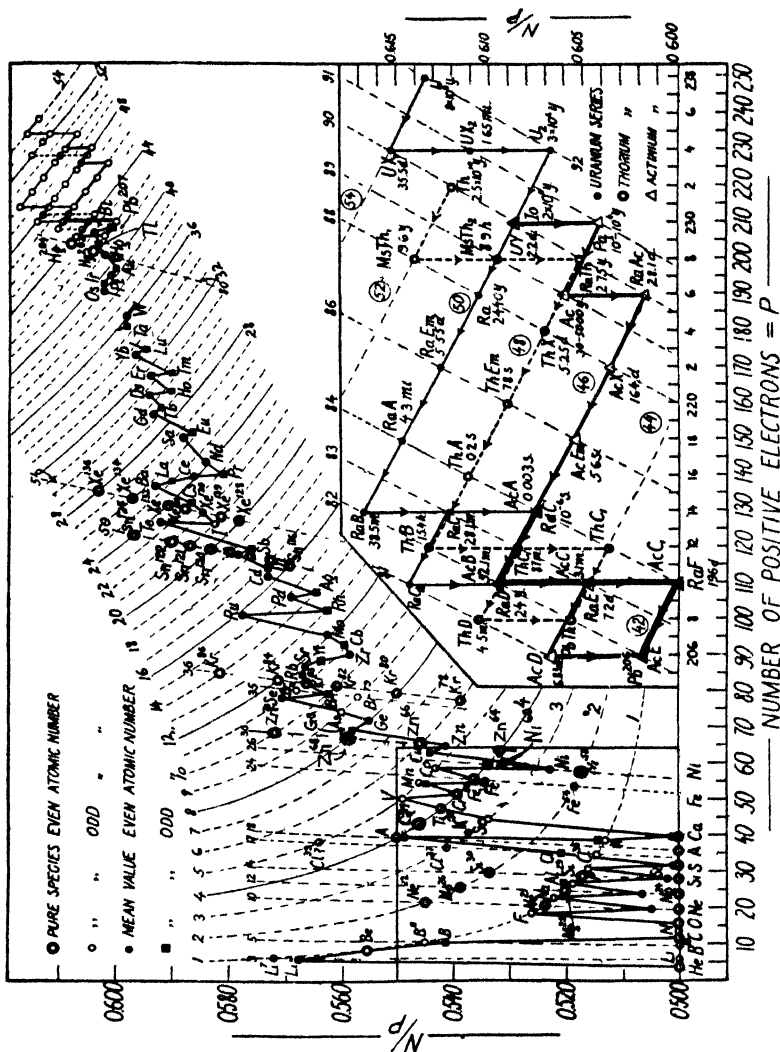


FIG. 8. RELATIVE NEGATIVENESS OF ATOM NUCLEI AS A FUNCTION OF THEIR TOTAL POSITIVENESS

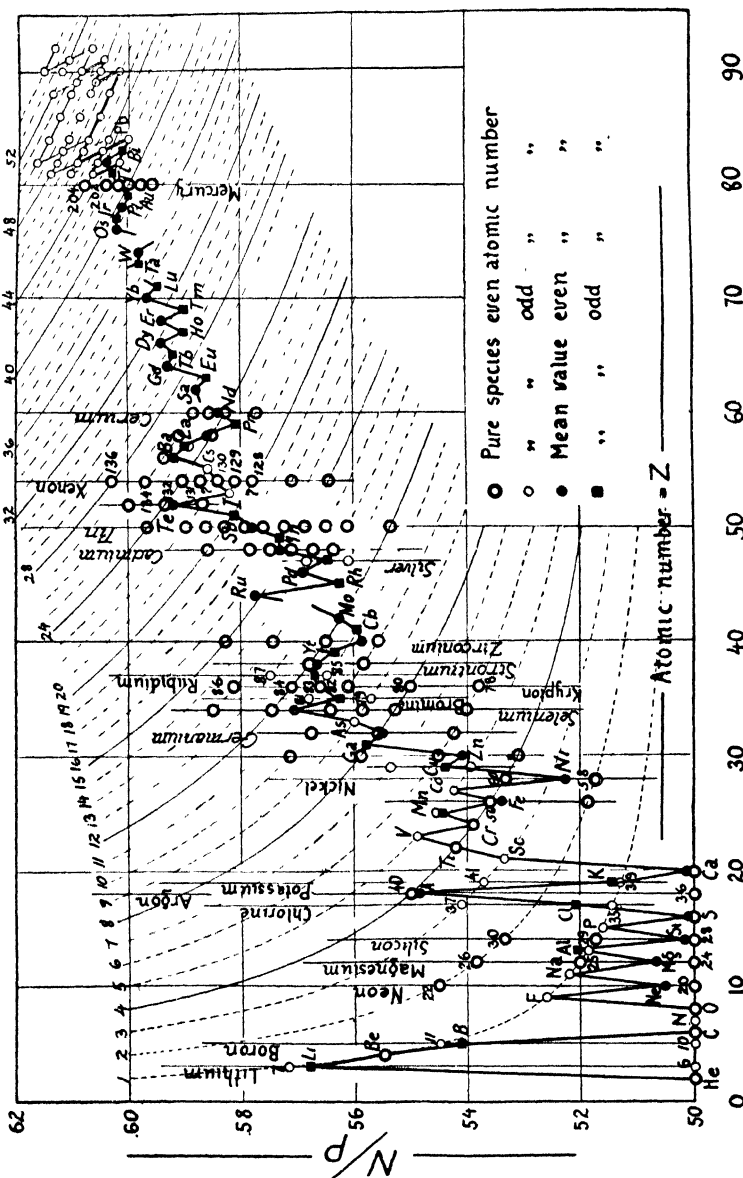


Fig. 9. RELATIVE NEGATIVENESS OF ATOM NUCLEI AS A FUNCTION OF THE NET POSITIVENESS

ratio $\frac{N}{P}$ is exactly 0.5, since such atoms constitute 80 per cent of the meteorites and 85 per cent of the earth's crust.

In figure 6 the electronic number is given on the Y-axis, and the protonic number on the X-axis. It is apparent that most of the light atom species up to atomic number 20 (Ca) lie either exactly on, or else only slightly above, the straight line of slope $1/2$. As the atomic number increases beyond 20 the curve, which represents the mean value for each element, rises farther from this line, not very far above it. The plot of the radioactive series indicates that all radioactive changes are represented by a shift along either a vertical line, or along a line whose slope is $1/2$.

The vertical extent of the band (fig. 6) in which atoms are stable, is very small. Thus for any definite value of the number (P) of protons in the nucleus, the range in the number of electrons (N) is very small for any atomic species (which is sufficiently abundant to have been discovered). Figure 7 gives the protonic number as a function of the atomic number.

The relation is shown in a more useful form in figure 8 which gives the ratio N/P on the Y-axis and, P, the number of protons on the X-axis. However, the plot shown in figure 9 is even more suggestive. It gives the atomic number or net nuclear positive charge ($Z = P - N$) on the X-axis, and the ratio N/P on the Y-axis. Thus the net positiveness of the nucleus is plotted against its relative negativeness.

2. General theory of nuclear stability

The most abundant light atom nuclei have atomic numbers and weights which may be represented by the formula a_J , in which a designates an α -particle, and J is any whole number (except 2) from 1 to 10 ($J = Z/2$). This shows either (1) that such nuclei are built up by the aggregation of α -particles, or (2) that, though not built up entirely by the aggregation of α -particles, the stability of nuclei whose composition is represented by that of a whole number of α -particles, is very much higher than that of other nuclei.

That a fast α -particle in some cases attaches itself to the nucleus of an atom has already been shown (4, 5, 6).

The heaviest known atomic species whose nucleus may be represented as α_J is calcium of atomic weight 40, for which J has the value 10, and for which N/P has the value $1/2$.

The preliminary assumption of the theory is that in atomic nuclei positively charged particles repel each other. However the binding action of negative electrons is so great that 3, 4, 5, 6, 7, 8, 9, or 10 α -particles will unite with each other in an atom-nucleus, even although each α -particle has a net positive charge of two. If somewhat more than 10 alpha particles are to unite, *two* additional (cementing) electrons must also enter the nucleus (1, 2, 7). Thus the charge of one α -particle does not appear in the net charge on the nucleus. If the number of α -particles rises above 16 (approximately) an additional pair of cementing electrons must be utilized, while in the heaviest atoms 13 such pairs of cementing electrons are present.

The theory may be stated in a much more general form as follows: *The net positive charge on the nucleus of an atom may rise as high as 20 while the relative negativeness (ratio of negative to positive electrons) remains at its most general value of 0.5, but in order that the nucleus may be stable and attain a higher net positive charge, the relative negativeness must increase to an extent which increases with increase in the net positive charge.* Thus the atom nucleus, in order to be stable, must have a relative negativeness (N/P) which matches its net positiveness (Z).

The band of stability is shown in figure 8. It is of great interest that the small number of atomic species contained in the rectangle in the lower left-hand corner of this figure constitute 99.9 per cent of all known material.

The usefulness of the relationship between stability and the relative negativeness (N/P) of the nucleus is illustrated by the fact that before the discovery of the isotopes of lithium four predictions were made concerning their existence. Three of these, made by others without a consideration of this relationship, were that the mass numbers are (1) 6 and 8, (2) 5 and 7, and (3) 6, 7, 8. Prediction (3), made by Rutherford, was the best

of the three, since its single error consisted in the assumption of the existence of a lithium isotope of isotopic number 2, with a value of N/P equal to 0.625. This is 0.125 in excess of the minimum value 0.5, while the highest value for any atom, even in the radioactive region, is only 0.117 in excess. Moreover, in the region of the plot in which lithium is found, the excess (0.072) for lithium, of mass number 7, is the maximum found for any atomic species. The prediction of an isotope of mass 5 is in discord with the fundamental relation, since the ratio N/P for such a species would be 0.4, or less than the minimum value.

The writer's prediction of lithium isotopes of masses 6 and 7 was made, on the basis of the relation between N/P and Z , in a paper presented to the American Chemical Society on April 12, 1920. The prediction was verified a year later by G. P. Thomson. A similar prediction, that boron would be found to have mass numbers 10 and 11, was also verified.

A study of the composition of the meteorites and of the crust of the earth, indicates that no atomic species of positive nuclear charge higher than 29, and also no species in which the ratio N/P is greater than 0.54 occurs in an amount greater than 1 atom in 10^4 (0.01 per cent).

While the ratio N/P is not the only factor upon which the stability of an atom depends, it is evidently of primary importance. Thus, beginning with species of low atomic number, it is found that lithium, beryllium, and boron, have ratios for the principal isotope of 0.572, 0.556, and 0.545. Now, all of these are abnormally high for such low atomic numbers (3, 4, and 5), and the species are found to be rare both on earth and in the meteorites. Carbon, the first species in the system (excluding the fundamental hydrogen and helium) in which the ratio falls to the normal value 0.5 is far more abundant, and is the lightest complex atom found in any considerable quantity in the meteorites. It is also about 26 times more abundant in the earth's crust than any one of the three other atomic species just mentioned. Nitrogen ($Z = 5$), though it has a value of N/P of 0.5, is not at all abundant. This may be explained as due to the fact

that it does not meet another extremely important stability condition discussed later.

The next three atomic species for which N/P is 0.5, are the principal isotopes of oxygen, magnesium, and silicon, with percentages of 61.94, 1.40, and 19.06, in the earth's crust, and 53.16, 9.86, and 13.82, in the meteorites. It is striking that these three atomic species constitute about 80 per cent of the atoms in known material. Neon is omitted from consideration since it is so light a gas that it escapes from the meteorites.

Until argon ($Z = 18$) is reached the most abundant isotope of each element of even atomic number, exhibits a value of 0.5 for N/P . With argon the ratio for the most abundant isotope rises to 0.55. This falls to 0.5 again for calcium ($Z = 20$), and rises again to 0.542 for titanium ($Z = 22$), 0.539 for chromium ($Z = 24$), and 0.536 for iron ($Z = 26$, $n = 4$). In general with increasing atomic number, the ratio N/P rises suddenly, then falls very gradually, and then repeats the procedure. The abundance of iron is remarkable if its relatively high atomic number (26) is taken into account, but this may be considered as due in part to a proper value of N/P (0.536) to give stability.

3. The relative negativeness (N/P) of the nucleus and the stability of radioactive atoms

The greater the ratio N/P , that is the relative negativeness of the nucleus, the more energetic is the emission of a β -particle which is negative, and the less energetic is the emission of an α -particle which is positive, provided the net positive charge on all of the nuclei compared is kept constant, that is when only isotopes are considered.

The above statement applies only to changes which actually occur. It puts into a new form an empirical relation discovered by Fajans.

The existence of this relation indicates that the term "stability of the nucleus" is not definite in its meaning since an atom nucleus which is extremely stable in the sense of an α -disintegration may be extremely unstable with respect to a β -disintegration.

V. GENERAL REPRESENTATION OF THE PERIODIC SYSTEM OF THE ATOMIC SPECIES

The periodic relations which concern atom nuclei are even more striking than those for the outer atom as exhibited in the periodic system of the elements as outlined by de Chancourtois, Newlands, Mendeléeff, and Mayer. The most general relation may be stated as follows: *Atomic stability, the abundance of atoms, and the number of atomic species of any type are all much higher for even than for odd numbers of* (1) *the electronic number* (N) (number of negative electrons in the nucleus), (2) *the protonic number* (P) (atomic weight or number of protons in the nucleus), (3) *the atomic number* (Z) (number of net positive charges on the nucleus of the atom, and (4) *the isotopic number* (n). *In addition there is a superimposed periodicity of 4 in the protonic number, and in the isotopic number.*

The new periodic system requires for its representation a figure in at least three dimensions. On account of the complexity of a 3 (or more) dimensional representation on the plane surface of a page, the relations of the system are exhibited in this paper by a series of projections.

The simplest two dimensional representation may be described as a double network of squares (fig. 10) with the atomic number on the X-axis, and the isotopic number on the Y-axis. The isotopic number for any atomic species other than hydrogen gives the number of the isotope according to a system in which the lowest possible isotope of any element is given the number *zero*. For example the isotopic numbers for the known species of a few elements are: He, 0; Li, 0 and 1; Be, 1; B, 0 and 1; C, 0; Mg, 0, 1, and 2; Al, 1; Si, 0, 1, and 2; Fe, 2 and 4. If the empirical formula of any nucleus is written $(p_+e)_Z (p_-e)_n$, then the subscript n gives the *isotopic number*, just as the subscript Z gives the atomic number.

The isotopic number is as important to distinguish the species as the atomic number to distinguish the element.

In figure 10 the *heavy coördinates represent even numbers and the light lines odd numbers*. Now it is remarkable that where heavy

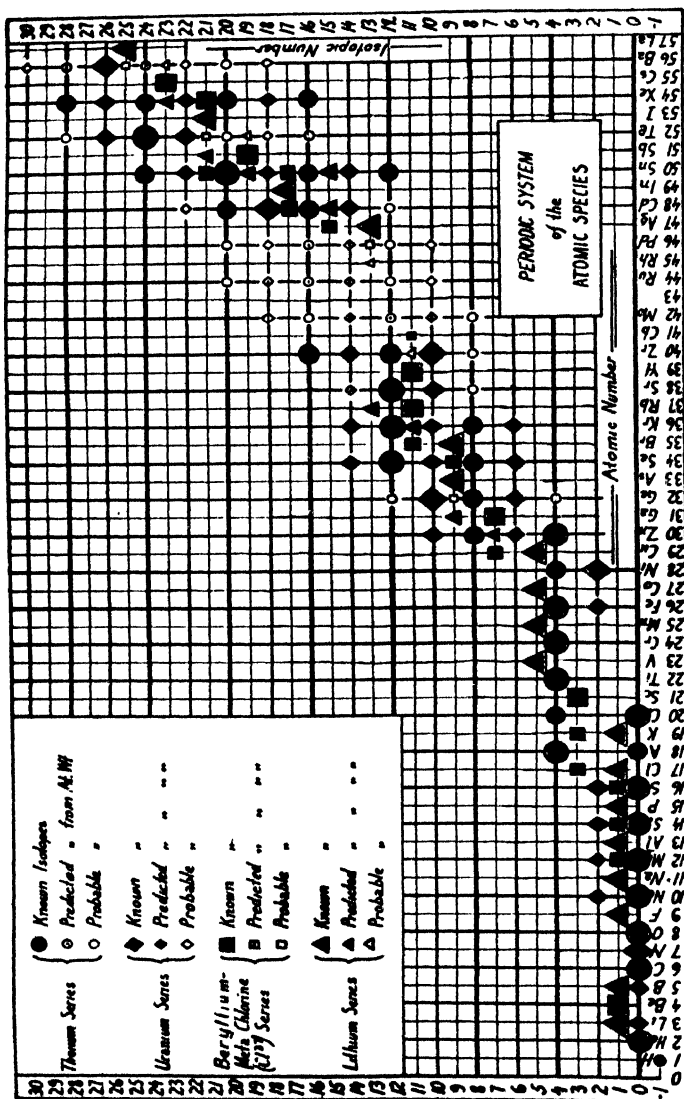


Fig. 10

lines meet heavy lines the greatest number of species occur (67 known species in the figure), and these are also by far the most abundant. Where light lines intersect light lines the number of known species is less (32), and the atoms of these species are in general far less abundant. Where heavy vertical lines intersect light horizontal lines the number of known species is again reduced by about 1/2 to 14 species. By far the fewest species occur where light vertical meet heavy horizontal lines. It is indeed remarkable that only 3 species (Li_0 , B_0 , and N_0) of this type are known in the whole system of 92 elements, and it is even more remarkable that all of the three lie on the *lowest heavy* line of the diagram, that of isotopic number 0.

The further details of these periodicities are discussed in the sections which follow.

VI. FIRST PERIODIC RELATION: HIGH STABILITY AND ABUNDANCE FOR ATOMS OF EVEN ATOMIC NUMBER (EVEN NUCLEAR CHARGE)

1. Introduction

The first periodic relation was discovered in the winter of 1914–1915 in an attempt to obtain evidence that the nuclei of complex atoms, that is all atoms other than hydrogen, are built up largely from α -particles, an idea suggested by the fact that the atomic weights of the light elements, with the single exception of glucinium, are divisible by 4 provided their atomic numbers are divisible by 2, that is the atomic weight and atomic number of helium. The development of the idea is shown best by a quotation from the second paper on the stability and abundance of the elements (2). It seems evident that the importance of the relation does not depend at all on the correctness of the point of view which led to its discovery.

It is the purpose of this paper to present a periodic system of the elements which relates their abundance to the structure of the nuclei of their atoms. This system is called *new* to distinguish it from the ordinary periodic system of Mendeleeff, to which it bears no relation. *It is important, too, that this is the first periodic relationship of the elements to be discovered which is independent of the ordinary periodic system.*

While the periodic variation of the abundance of the elements as here presented might have been discovered empirically, the fact is that the relations found were first predicted from the standpoint of theory, and were afterward found to be true. It therefore seems important to trace, at least in part, the reasoning by which this new law was discovered, even although its validity in no way depends upon the truth of the hypotheses which were used. However, the bearing of these new facts upon the important problem of the evolution of the elements, can best be understood if the subject is approached from a theoretical standpoint.

In the fifth paper of this series evidence has been presented for the theory that the variation in the chemical, and in such physical properties of the elements as cohesion, atomic volume, compressibility, coefficient of expansion, melting point, etc., depends first of all upon the arrangement in *space* of the negative electrons in the atom external to the nucleus, and also upon the number of such electrons. The number of these electrons presumably depends upon the nuclear charge. The structure of the nucleus determines what is called the stability of the atom, since it is not considered that the atom breaks up unless the nucleus disintegrates. For example, when negative electrons are given off under the influence of light, or from other similar causes, the atom is said not to decompose, but only to become electrically charged (ionized).

Now, the earlier papers have shown that the 91 elements other than hydrogen, of our ordinary system, fall into two series. At least among the elements of lower atomic weight, the atoms which have *even atomic numbers* are in general built up from helium atoms, and therefore may be said to have the general formula $n\text{He}'$, where the prime is added to indicate that these elements are intra-atomic, not chemical, compounds. The odd-numbered elements, beginning with lithium, seem in general to have the formula $n\text{He}' + \text{H}_3'$. Thus these elements fall into two series which may be distinguished as *even* or *odd*.

Now, if the variations in the chemical and the ordinary physical properties of the cohesive type (properties of aggregation) are dependent upon the arrangement and number of the electrons external to the nucleus, which seems probable if the nucleus is extremely minute, as the results obtained by Rutherford and his students⁶ seem to indicate, then the differences between the odd- and even-numbered atoms with

⁶ Their results were obtained by the scintillation method. The photographic method, as applied by Harkins, Ryan, Shadduck, and Shah shows the apparent small size of the nucleus even more simply.

respect to nuclear structure would not be expected to be apparent in these properties. That this expectation is confirmed is easily seen, since neither with respect to chemical nor physical properties is there any difference apparent which shows a variation in periods of two elements from even to odd, or odd to even.

2. *The hydrogen-helium system and the abundance of the elements*

It might be expected, however, that the composition of the nucleus should affect its own stability, which from radioactive evidence means the stability of the atom. From this standpoint it might be reasonable to suppose that the atoms of one of the series, the even or the odd, should be more stable than those of the other. Now, unfortunately, there is no known method of testing the stability of the lighter atoms, but it might seem, at least at first thought, that the more stable atoms should be the more abundantly formed, and to a certain extent this is undoubtedly true. If then, at the stage of evolution represented by the solar system, or by the earth, it is found that the even-numbered elements are more abundant than the odd, as seems to be the case, then it might be assumed that the even-numbered elements are on the whole the more stable. However, there is at least one other factor than stability which must be considered in this connection. The formula of the even-numbered (light) elements has been shown to be $n\text{He}'$. Now, since the formula for the odd-numbered elements is $n\text{He}' + \text{H}_3'$, it is evident that, if the supply of H_3' was relatively small at the time of their formation, not so much material would go into this system. This would be true whether the H_3' represents three atoms of hydrogen or one atom of some other element.⁶

3. *The composition of meteorites as related to the structure of complex atoms*

There is, however, material available of which accurate quantitative analyses can be made, and which falls upon the earth's surface from space. The bodies which fall are called meteorites, and no matter what theory of their origin is adopted, it is evident that this material comes from much more varied sources than the rocks on the surface of the earth.

⁶ If it were one element it might have an atomic number of 0, 1, or 2, and thus be a neutron, an atom of hydrogen, or an atom of helium, of mass 3. The stability relations are not favorable to the existence of such a unit as a stable atom.

In any event, it seems probable that the meteorites represent more accurately the average composition of material at the stage of evolution corresponding to the earth than does the very limited part of the earth's material to which we have access. At least it might seem proper to assume that the meteorites would not exhibit any special fondness for the even-numbered elements in comparison with the odd, or *vice versa*, any more than the earth or the sun as a whole, at least not unless there is an important difference between these two systems of elements, which is just what it is desired to prove.

A preliminary study of the *most recent analyses* of meteorites of different classes showed that, either for any one class or for the meteorites as a whole, *the even-numbered or helium system elements are very much more abundant than those of the old-numbered or helium-hydrogen system*. For a more detailed study use was made of the older but much more complete and more valuable data collected by Merrill and by Farrington (10), who suggest that the average composition of meteorites may represent the composition of the earth as a whole.

The results obtained by averaging the analysis of 318 iron and 125 stone meteorites, 443 in all, show that the first seven elements in order of abundance are iron, oxygen, nickel, silicon, magnesium, sulfur, and calcium; and *not only do all of these elements have even atomic numbers, but in addition they make up 98.8 per cent of the material of the meteorites*. Of the remaining elements present to a great enough extent to have an appreciable effect upon the percentage values, 7 are odd and 5 are even, but in all only 1.22 per cent are odd numbered, while 98.78 per cent are even. Of the iron meteorites, 99.22 per cent of the material is made up of even-numbered elements, and of the stone meteorites, 97.59 per cent.

The percentages by weight given above are proportional to the number of hydrogen atoms utilized in building up the respective atoms. The general relations are not at all affected if the atomic percentages (number of atoms per 100 present) are used. The six most abundant elements on this basis all have even atomic numbers, and in order of abundance are iron, oxygen, silicon, magnesium, nickel, and sulfur.

In studying the relative abundance of the elements the ideal method would be to sample one or more solar systems at the desired stage of evolution, and to make a quantitative analysis for all of the 92 elements of the ordinary system. Since this is impossible, even in case of the

earth, it might be considered that sufficiently good data could be obtained from the earth's crust, or the lithosphere.

However, there are several important factors which cause our knowledge of the quantitative composition of the earth's crust to be of much less value for the solution of our problem than it might seem to possess on first thought. In the first place the quantitative analyses which have been made represent the composition of only the mere skin of the earth, the depth of which does not exceed the ten to twenty miles caused by geologic displacements. The surface of the earth has been markedly influenced both by igneous processes which have resulted in magmatic differentiation, and by weathering, solution, and redeposition. For example, the common idea that sodium is a very abundant element undoubtedly has its origin in the fact that the solubility of its salts has caused their very considerable concentration in the oceans. Again, the fact that the salts of sodium are much more fusible than similar salts of the alkaline earths and most other metals in the rocks, has probably caused it to be segregated by magmatic solution and redeposition. Thus, while in the average igneous rock found on the surface of the earth there seems to be about 2.23 per cent of sodium, it is not improbable that this is a larger percentage than would be found if the whole material of the earth could be taken as a sample for analysis.

If the sun is next considered it is found that although a large amount of its surface is exposed to us for spectroscopic investigation, the spectroscope gives no accurate measure of the quantitative composition, and that its findings are largely influenced by the height in the gaseous envelope of the sun at which the observation is taken.

Table 4 in columns 10 to 13, gives the average composition of iron and stone meteorites, arranged according to the atomic numbers. It will be noted that the *even-numbered elements are in every case more abundant than the adjacent odd-numbered elements*. The helium group elements form no chemical compounds, and are all gases, so they could probably not remain in large quantities in meteorites. For this reason, and also because the data are not available, the helium or zero group is omitted from the table. The only criticism which could be made of the system of averaging used in these columns, which is that of including all accurate analyses, is that it places undue emphasis upon the iron as compared with the stone meteorites. However, since the two relations shown are true for each class of meteorites separately, it is evident that they will be true whatever system of averaging may be chosen.

The results given in the last two columns of table 4 are expressed

22	Titanium	...	0 01	0 108	Trace	0 005	...	0 01	...	0 005	0 005	...	
23	Vanadium	0 321	Trace	0 13	...	0 09	...	0 07	...	0 01	0 13	...	
24	Chromium	...	0 28	...	0 58	...	0 224	...	0 06	...	0 03	0 06	
25	Manganese	0 14	0 224	...	10 57	52 93	...	90 64	12 79	...	
26	Iron	...	24 32	...	24 32	0 02	0 04	
27	Cobalt	0 05	0 017	...	0 54	...	0 44	...	0 31	...	0 59	...	
28	Nickel	...	1 31	...	1 26	6 50	4 52	...	8 50	...	0 76	
29	Copper	0 01	0 014	...	0 005	0 005	...	0 02	...	
Total		...	97.59	2.41	97.06	2.94	2.978	97.89	2.11	98.78	1.22	98.57	1.425	99.22	0.78	97.95	
			100%			100%			100%			100%			100%		

* 40, zirconium, none; 50, tin, none; 56, barium, none; elements 44, ruthenium, 46, palladium; 77, iridium; and 78, platinum, were found in stone meteorites, while 30, zinc; 33, arsenic; 51, antimony, 50, tin; 79, gold; 74, tungsten; and 92, uranium, were not found in the investigations reported by Merrill.

† It has been found by Farrington that the known falls of meteorites are in the ratio of 350 stone meteorites to 10 of iron.

graphically in figure 11, where the atomic number are given as abscissae, and the percentage abundance of the elements is plotted on the Y axis. This figure shows that there is a very marked periodicity in the abundance of the elements in periods of two elements each. The peaks which represent the even-numbered elements are shown to be extremely high in comparison with the troughs in which the odd-numbered elements lie

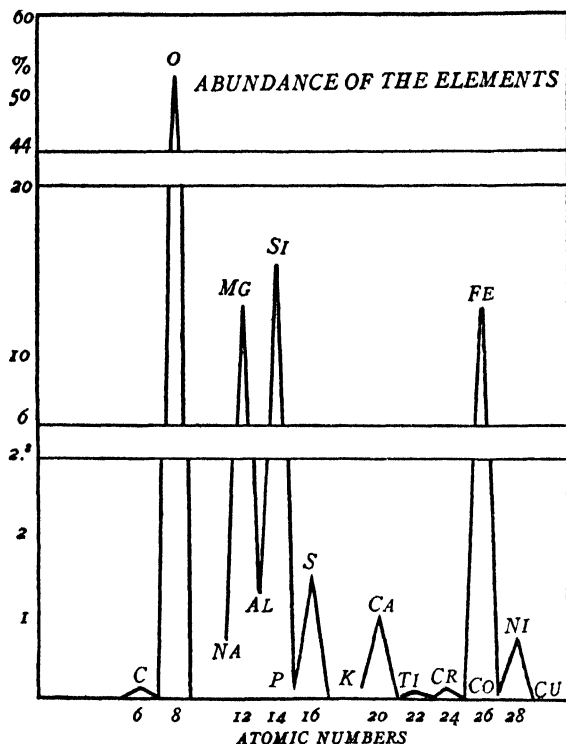


FIG. 11. ABUNDANCE OF THE CHEMICAL ELEMENTS IN THE METEORITES (ATOMIC PERCENTAGE)

It is interesting to note that where two high peaks for even-numbered elements lie adjacent, between them the abundance of the odd-numbered element also becomes high in comparison with that of the other odd-numbered elements. Thus, by far the most abundant of the odd-numbered elements, aluminium and cobalt, lie in each case between two extremely abundant even-numbered elements.

This would seem to indicate that there is some relationship between the even and the odd series, and that the abundance not only follows the periods of two elements but also larger periods or waves, which cannot be said to have regularity in length and do not correspond with any of the divisions found in the ordinary periodic system.

4. A comparison of stone and iron meteorites with respect to their composition

Table 4 gives the average composition, in terms of both atomic and weight percentage, of both the iron and stone meteorites. This table shows in a very marked way that the extreme variation

TABLE 5

Abundance (percentage by weight) of the elements in the lithosphere, illustrating the predominance of elements of even atomic number

ATOMIC NUMBER	ELEMENT	EVEN	ODD
8	Oxygen	47.33	7.85
14	Silicon	27.74	
13	Aluminium		
26	Iron	4.50	
20	Calcium	3.47	2.46
11	Sodium		
19	Potassium		
12	Magnesium	2.24	
22	Titanium	0.46	
		85.74	12.77

in composition between the iron and the stone meteorites does not affect markedly the predominance of the even-numbered elements. When calculated from the standpoint of the atomic percentages the iron meteorites contain less than 1 per cent of odd-numbered elements, and the stone meteorites less than 3 per cent.

It is very remarkable that in none of the average analyses presented does the percentage of any odd-numbered element, whether calculated by weight or by atomic proportions, rise higher than 1.53 per cent, which is the highest value given by Merrill for aluminium in any of his averages, while among the even-numbered elements large percentages are common, and range as high as 90.64 per cent.

5. The predominance of elements of even atomic number as illustrated by the composition of the earth's crust

Since it has been shown that the elements of even atomic number are 70 times as abundant in the meteorites as those of odd atomic number, it should be of interest to see if the relations on the surface of the earth are at all similar. It should be remembered, however, that the crust of the earth is much more local in its character, and cannot therefore be expected to give such good evidence as that presented by the meteorites. It is found, however, that the elements of even number again show their predominance, since of the five most abundant elements, four have even atomic numbers, while of the first nine, six are even numbered, according to Clark's average for the composition of the lithosphere, as presented in table 5.

The six even-numbered elements make up 85.74 per cent of the lithosphere, while the three odd-numbered elements amount to only 12.77 per cent.

Table 6 gives another set of data for all of the elements as recalculated from a recent estimate made by Clarke and Washington.

The latter table shows that in the twenty top miles of the earth's crust the atoms of elements of even atomic number are about nine times as numerous as those of odd atomic number.

Since the data presented for the composition of the lithosphere suffer from the disadvantage for the present purpose, that the surface of the earth has been subjected to extensive processes of differentiation, it would be of extreme value if evidence could be obtained in regard to elements which have not been thus affected. While this cannot be done, it would seem that elements which are very much alike both chemically and physically should be affected more nearly to the same extent than those which differ widely. Now the rare earths are so similar that they are separated in the laboratory only with great difficulty, so the writer first made an estimate of the relative abundance of these elements, and then secured estimates made independently by two noted authorities, Profs. C. James and C. W. Balke. It is of interest

TABLE 6
*Abundance of the elements in the crust of the earth**
 (Atomic percentage for a twenty-mile crust)

ORDER	ELEMENT	ATOMIC NUMBER	HYDRO- GEN	LIGHT ELEMENTS		HEAVY ELEMENTS	
				Even	Odd	Even	Odd
1	Oxygen	8	9.48	57.44			
2	Silicon	14		18.23			
3	Hydrogen	1					
4	Aluminium	13			5.55		
5	Sodium	11			2.27		
6	Iron	26		1.68			
7	Calcium	20		1.68			
8	Magnesium	14		1.59			
9	Potassium	19			1.223		
10	Titanium	22		0.273			
11	Carbon	6		0.152			
12	Phosphorus	15			0.093		
13	Chlorine	17			0.088		
14	Sulfur	16		0.052			
15	Manganese	25			0.041		
16	Fluorine	9			0.030		
17	Chromium	24		0.024			
18	Nitrogen	7			0.022		
19	Vanadium	23			0.015		
20	Lithium	3			0.014		
21	Zirconium	40				0.012	
22	Barium	56				0.011	
23	Nickel	28		0.010			
24	Strontium	38				0.007	
25	Ce-Yt						(0.0037)†
26	Copper	29					0.0033
27	Glucinium	4		0.0019			
28	Boron	5			0.0017		
29	Zinc	30				0.0012	
30	Cobalt	27			0.0010		
31	Lead	82				0.0002	
Total			9.48	81.1329	9.3487	0.0314	0.0033 (0.0037)† = 100.00

* Recalculated from a recent estimate by Clarke and Washington.

† In the analyses for these elements those of odd and those of even atomic number are reported together.

TABLE 6—*Continued*

0.000,0×	0 000,00×	0 000,000×	0 000,000,0×	0 000,000,00×
Rubidium	Cadmium	Bismuth	Platinum	Thallium
Arsenic	Mercury	Tungsten	Tellurium	0 000,000,000 ×
Molybdenum	Iodine	Thorium	Gold	Indium
Tin	Caesium	Cb-Ta		Gallium
Bromine	Antimony	Uranium		0.000,000,000,0×
Scandium		Silver		Germanium
		Selenium		0 000,000,000,000 ×
				Radium

that these three estimates were all in agreement. They are given in table 7, in which c indicates common in comparison with the adjacent elements and r represents rare; ccc represents a relatively very common element, etc. The comparison is only a rough one, but it indicates that the even-numbered rare earths are much more abundant than the adjacent ones which are odd.

TABLE 7

The predominance of even-numbered elements among the rare earths

ATOMIC NUMBER	ABUNDANCE	ELEMENT	ATOMIC NUMBER	ABUNDANCE	ELEMENT
55	c	Caesium	63	rr	Europium
56	ccc	Barium	64	r	Gadolinium
57	c	Lanthanum	65	rrr	Terbium
58	cc	Cerium	66	r	Dysprosium
59	r	Praseodymium	67	rrr	Holmium
60	c	Neodymium	68	r	Erbium
61	rrr	Illinium	69	rr	Thulium
62	c	Samarium			

A striking confirmation of the idea that each element of even atomic number is more abundant than the two adjacent elements of odd number was obtained in an extremely extensive investigation of the composition of rare earth minerals by Goldschmidt and Thomassen (11). Their quantitative results confirm the relation in an extremely remarkable way, as is shown in figure 12.

The above results may be summarized in the statement that

in the formation of the elements much more material has gone into the elements of even atomic number than into those which are odd.

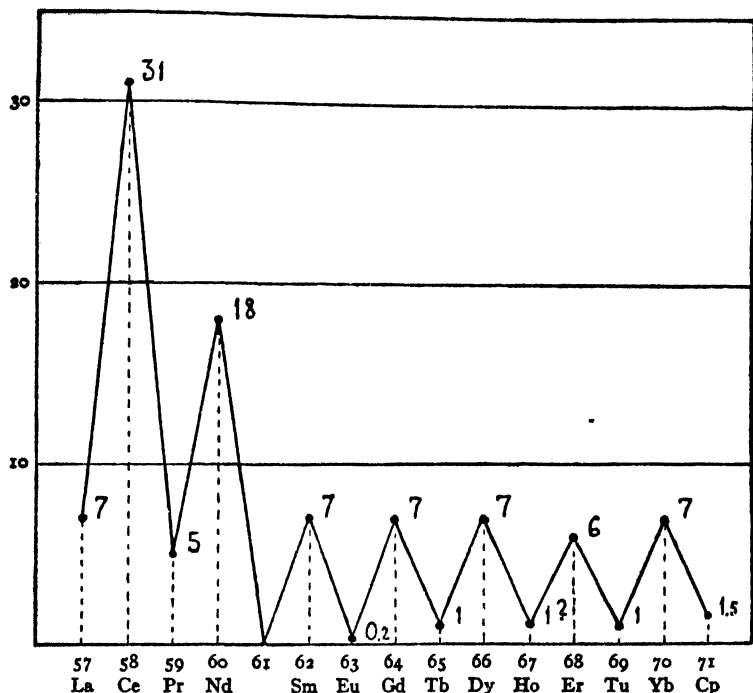


FIG. 12. PERIODICITY OF TWO IN THE ABUNDANCE OF THE RARE EARTHS AS A FUNCTION OF THE NUMBER, AS DETERMINED BY GOLDSCHMIDT AND THOMASSEN. THE HEIGHT OF EACH POINT REPRESENTS THE ABUNDANCE OF THE RARE EARTH ELEMENT IN MINERALS ON THE SURFACE OF THE EARTH

Obviously each element of even number (number given below) is very much more abundant than either adjacent element of odd number. This accords exactly with the predictions of the hydrogen-helium theory of atom building developed eight years before the work represented above was done by Goldschmidt and Thomassen.

6. The unknown elements

Of the ninety-two elements within the limits of our present system, all but two have been discovered. It is of interest to note that *both of the unknown elements have odd atomic numbers*, and that they lie in the radioactive region.

7. *The greater stability of radioactive elements of even number*

Among the radioactive elements it will be seen too that if the most stable isotope of each element is considered, then in each case the odd-numbered element has either a shorter period than the adjacent even-numbered elements, or else is entirely unknown. This is shown by table 8.

8. *Classification into light elements (atomic numbers 1 to 27) and heavy elements (atomic numbers 28 to 92)*

It has been seen that an important natural classification of the elements with respect to their nuclear characteristics is into those

TABLE 8
Half life of the most stable isotopes of the radioactive elements

ATOMIC NUMBER	ELEMENT	ONE-HALF PERIOD OF MOST STABLE ISOTOPE	
		Even	Odd
92	Uranium	5 billion years	1.15 minutes
91	Uranium X ₂		
90	Thorium	18 billion years	
89	Actinium		Period unknown, but almost certainly less than radium
88	Radium	1730 years	Element undiscovered
87			
86	Niton	3 85 days	Element undiscovered
85			
84	Polonium	136 days	

of even and those of odd atomic number. A second important natural classification, introduced in 1915 (1, 12), is into light elements up to atomic number 27, and heavy elements between atomic numbers 28 and 92. Of the latter elements 81 to 92 may be classed as specifically radioactive.

9. *Primary and secondary factors which determine the abundance of the elements*

It is evident from the preceding discussion that there is a direct relation between the abundance of the elements and the atomic

number, and that there is no very marked relation to the periodic system of Mendeléeff. This indicates that the abundance of the elements is determined in a primary sense by the stability relations of the complex atoms themselves, and also by the stability of any electron-proton groups, such as the alpha particle, from which they are built. However, it is obvious, *even if in the formation of the elements there is no relationship to the ordinary periodic system, that in all of the differentiative processes which occur after the original formation, the chemical and physical properties would play their part, so that if any special material is taken for consideration, relationships to that system would appear more and more as the process of differentiation takes place.*

The fact that in the meteorites the elements of even atomic number are in all cases more abundant than the adjacent elements of odd number, seems to indicate that in such bodies the process of differentiation has not proceeded to such an extent as to cause the chemical and physical properties to have an influence sufficient to obscure in any marked degree the most general of the abundance relations due to nuclear stability, but it does not show that there has been an entire absence of differentiation. The composition of the earth's surface, as compared with that of the meteorites, shows a less rigid, but very marked adherence to the odd and even system, but certain groups of elements related in chemical and physical properties, are relatively more abundant. Thus the alkali metals, sodium and potassium, the halogens, and aluminium are present in much larger quantities. In the meteorites calcium is much less abundant than magnesium, while on earth both elements are of the same order of plentifulness.

10. The number of isotopes per element

It is evident that if the elements of even atomic number have in general more highly stable atoms than those of odd number, they should also have more kinds of atoms, that is a greater number of isotopes. The fact that this is true for the radioactive elements was used by the writer as one basis for the development of the relation, and was noted independently by N. F. Hall (13), to whom the idea was suggested by the relation as published.

The average number of isotopes of the known **radioactive** series averages 4.5 for elements of even, and 1.8 for those of odd atomic number.

That the periodicity is extremely striking in the case of ordinary elements is shown by figure 13 which gives the number of known isotopes for elements 29 to 40 and 47 to 55. This

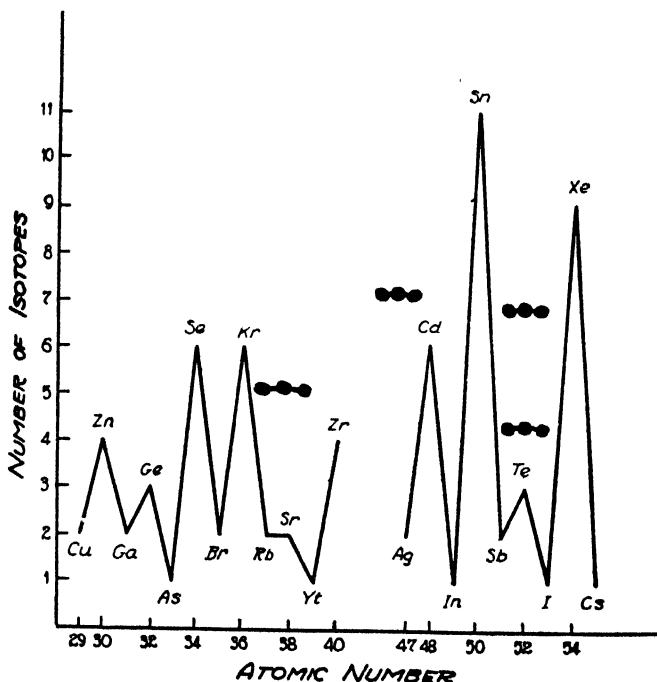


FIG. 13. NUMBER OF ISOTOPES PER ELEMENT

periodicity is much less pronounced in the region defined by the writer as that of the "light elements," that is in the region of atomic numbers from 2 to 28.

VII. SECOND PERIODIC RELATION: HIGH STABILITY AND ABUNDANCE FOR ATOMIC SPECIES OF EVEN ELECTRONIC NUMBER

The second periodic relation is even more general and important than the first.

The number N which is defined as the difference between the protonic number P , and the atomic number Z , may be designated as the *electronic number*.

$$N = P - Z$$

Since P and Z are experimental quantities, N is also an experimental quantity. From the theoretical point of view N gives

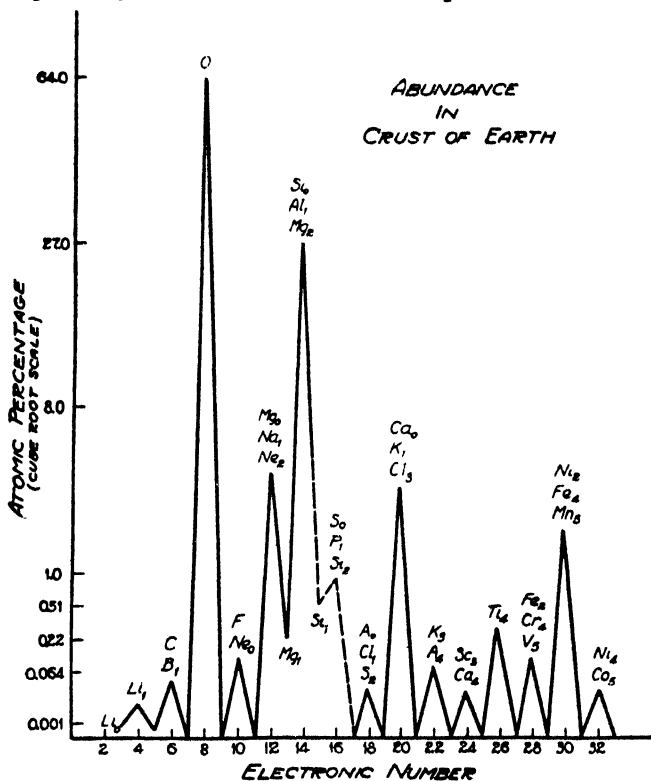


FIG. 14. ABUNDANCE OF ATOMIC SPECIES FOR ELECTRONIC NUMBERS 3 TO 33

the number of negative electrons in the nucleus of the atom, and is a whole number.

In general atoms of even electronic number are extremely more abundant and stable than those of odd number, and the number of stable species of atoms is also very much larger. That high abun-

dance is associated with even electronic number is exhibited in a remarkable way by figures 14 and 15 which represent the abundance in the earth's crust and in the meteorites respectively. The extraordinary nature of the contrast in abundance will not be realized unless it is noted that the *cube root* of the atomic per-

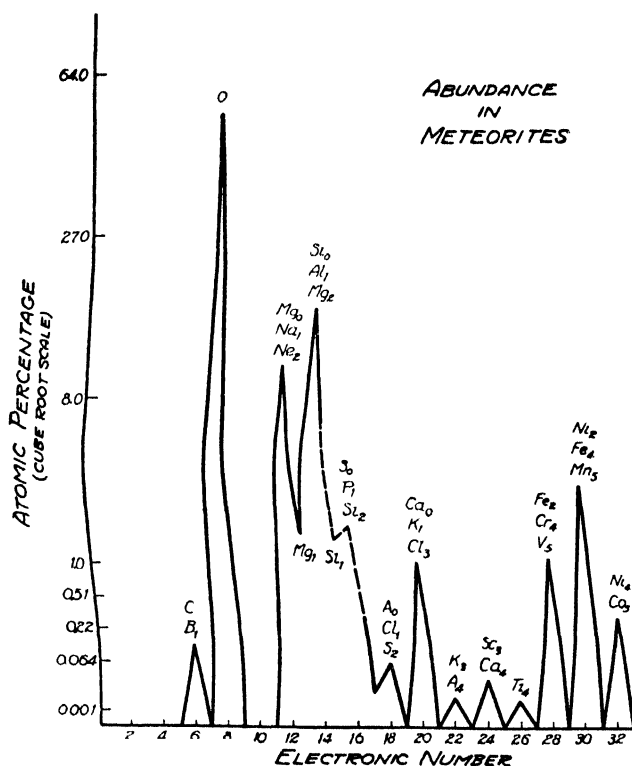


FIG. 15. ABUNDANCE OF ATOMIC SPECIES FOR EACH ELECTRONIC NUMBER

centage is plotted on the Y-axis, and the effect of the cube root plot is to obscure the great contrast in abundance. Thus the peaks are much higher in comparison with the troughs than they appear to be.

In general the abundance of the atomic species for any even electronic number is much greater than for the two adjacent odd

electronic numbers. If the percentages themselves were to be plotted it would be necessary to make figure 14 more than 100 meters high in order to give the points for electronic numbers 3 and 5 the height given to them in the figure.

That the periodicity of 2 in the number of atomic species found for the atomic number appears also in the electronic number is exhibited by figure 16 also, since it shows that between electronic numbers 25 and 51 the number of known species for each even electronic number is much higher than for the two adjacent odd

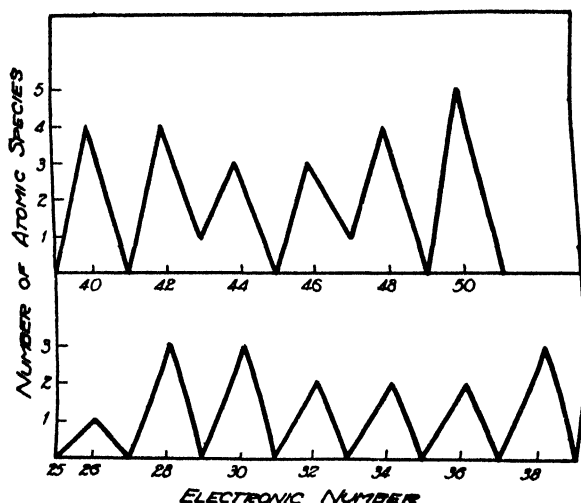


FIG. 16. PERIODICITY OF TWO IN THE NUMBER OF ATOMIC SPECIES AS A FUNCTION OF THE ELECTRONIC NUMBER

numbers. Of the 112 atomic species thus far found between atomic numbers 2 and 60, only 16 have an odd, while 96 have an even electronic number.

Among the radioactive elements all of the most stable species are those of even electronic number: for example, the half-life of U is 5 billion years; of U_2 , 2 million years; of Io, 200 thousand years, and of radium 1730 years. The longest half-life for a species of odd electronic number is only 5 days (RaE).

Also, among the ordinary elements the most abundant isotope

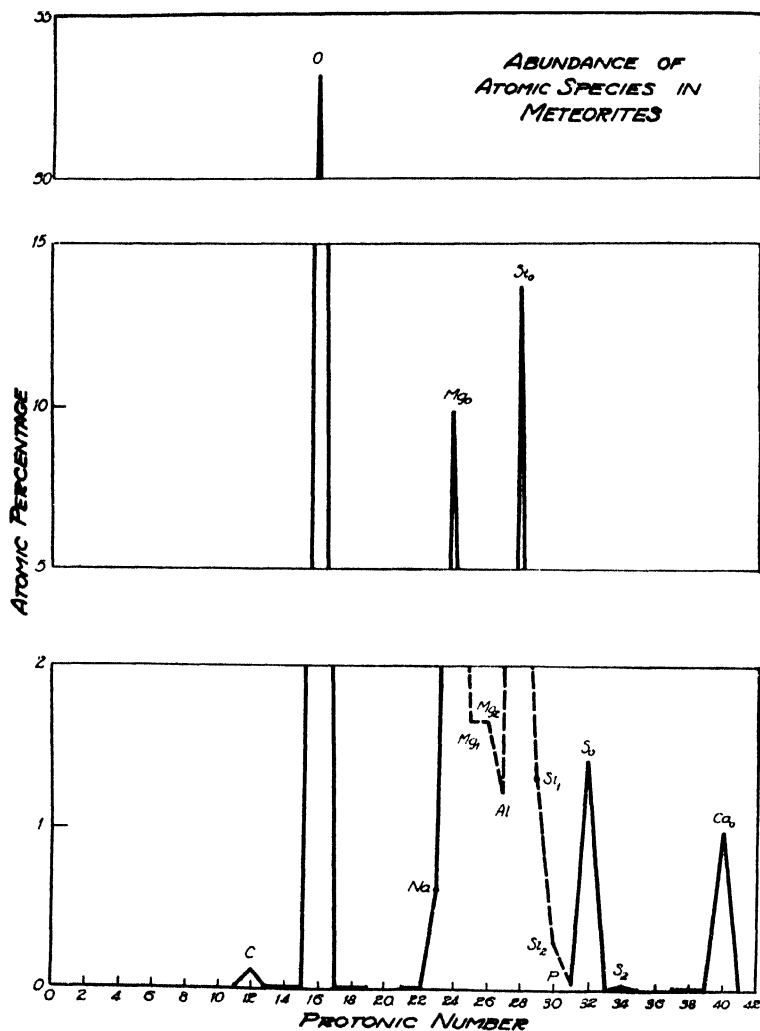


FIG. 17. ABUNDANCE OF ATOMIC SPECIES FOR EACH PROTONIC NUMBER (ATOMIC WEIGHT)

of every element thus far investigated (atomic numbers 2 to 40, and 47 to 60) has an even electronic number, with the two exceptions of nitrogen and xenon.

VIII. THIRD PERIODIC RELATION: HIGH STABILITY AND ABUNDANCE FOR ATOMS WHOSE PROTONIC NUMBER (ATOMIC WEIGHT) IS DIVISIBLE BY 4, AND A SECONDARY PERIODICITY OF 2

The abundance of the light atomic species exhibits an extremely marked periodicity of 4 with respect to the protonic number.

TABLE 9

Atomic percentage of atomic species in meteorites, showing the importance of the weight 4

ELEMENT	ATOMIC WEIGHT						
	4q	3q	4q not 3q	3q not 4q	4q - 3	4q - 2	4q - 1
Carbon	0.12	0.12					
Oxygen	53.16		53.16				
Sodium	0.62		0.62		0.62		
Magnesium.	9.86	9.86				1.65	1.65
Aluminium		1.21		1.21	1.21		
Silicon ²²	13.82		13.82				1.38
Phosphorus					0.06		
Sulphur.	1.46		1.46				
Potassium		0.11		0.11	0.11		
Calcium	0.97		0.97				
Titanium ...	0.005	0.005					
Chromium.....	0.13		0.13				
Manganese.					0.06		
Iron ⁴⁴	12.30		12.30				
Cobalt					0.04		
Totals ...	92.45		82.46	1.32	2.10	1.65	3.03

q = a whole number.

In figure 17, which refers to the meteorites every high peak in abundance occurs for species whose atomic weight is divisible by 4. Thus such peaks occur for oxygen, silicon ($n = 0$), magnesium ($n = 0$), sulfur ($n = 0$), calcium ($n = 0$), and carbon.

Atoms in which the protonic number is even are much more abundant than those in which the number is odd; about 18

times as abundant in the meteorites and $6\frac{1}{2}$ times as abundant in the crust of the earth, but this contrast is due mainly to the periodicity of 4, and only to a slight extent to that of 2.

The number of known species of even protonic number is found to be 71, and of odd protonic number 49, if the summation is made between protonic numbers 4 and 89, and between 109 and 146, that is in the regions most carefully investigated. Thus the emphasis upon even numbers is not at all as striking as if electronic numbers are involved.

The most abundant isotope of each element of even number from 4 to 28 has an atomic weight divisible by 4. Between atomic numbers 28 and 60, however, the atomic weight is divisible

TABLE 10

Classification of atomic species according to even and odd number of electrons

	ABUNDANCE IN ATOMIC PERCENTAGE	
	Earth's crust	Meteorites
Class I, N = even, P = even	87 4	95 4
Class II, N = even, P = odd	10 8	2.1
Class III, N = odd, P = odd	1.8	2 5
Class IV, N = odd, P = even	0 0007	0 0

by 4 in 8 known cases, by 2 but not by 4, in 5 cases. That is the periodicity of 4 becomes less general for even atomic numbers greater than 26, and is changed largely to a periodicity of 2.

The extreme emphasis on the weight 4 (weight of the alpha-particle) is exhibited in table 9, in which q is a whole number.

The totals of the columns in a similar table for the earth's crust are almost the same, except that the column $4q - 3$ adds up about 8 atomic per cent, and therefore is much more prominent. In both tables the abundance of the atomic species whose atomic weights are divisible by 4 is much greater than that of those divisible by 3, and the latter seems to gain its relative importance only through the inclusion of those weights which are common to the column $4q$.

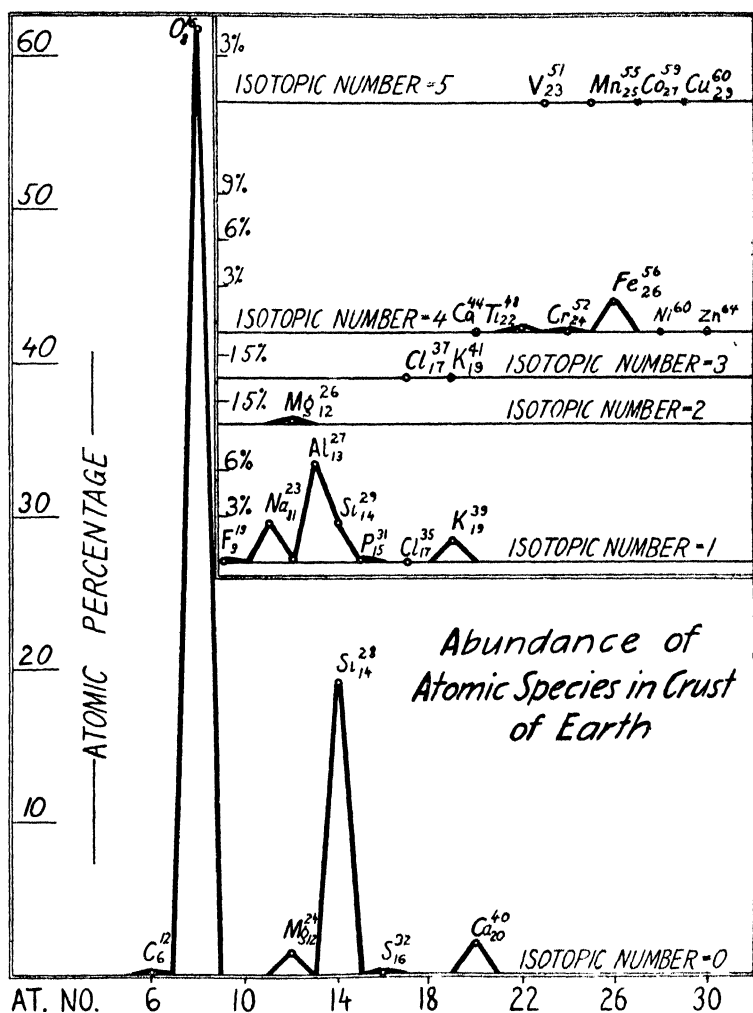


FIG 18

IX. CLASSIFICATION OF ATOMIC SPECIES (ISOTOPES) ACCORDING TO EVEN AND ODD NUMBER OF ELECTRONS AND PROTONS (15)

The atomic species represented in figure 10 may be classified in a perfectly natural way according to the type of line intercepts

on which they lie in the figure, but more fundamentally according to the evenness or oddness of the number of electrons and protons present in the nucleus. The classification is given in table 10.

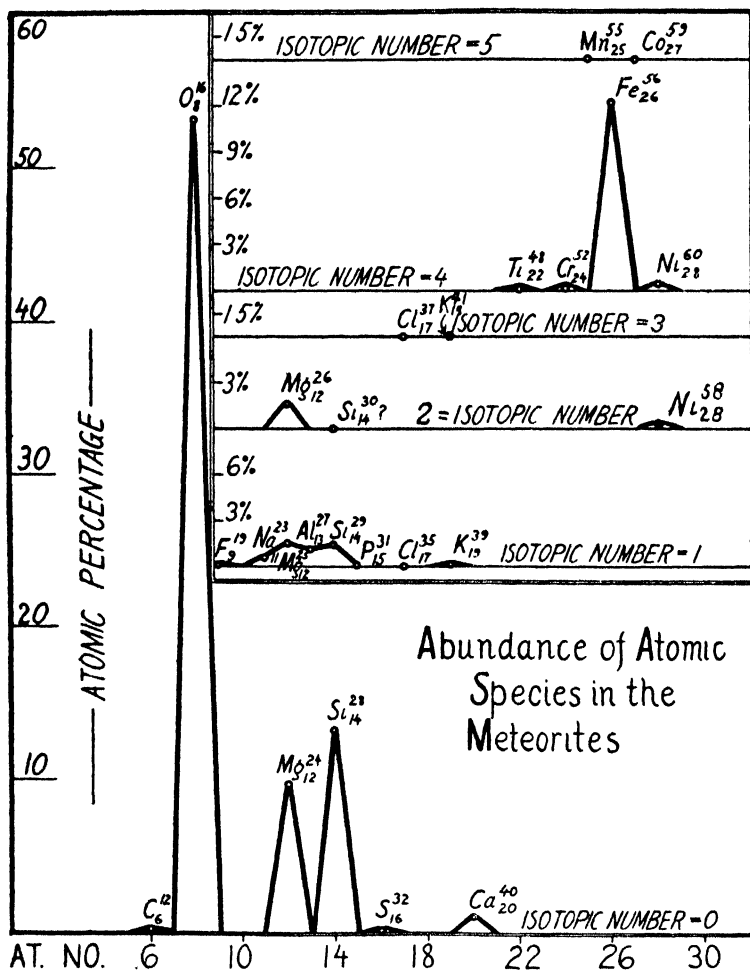


FIG. 19

It is apparent that N is usually even whether P is even or odd, but that if N is odd, then P is also almost always odd. Thus

there is a certain matching of P with N. The class which is most rare, Class IV, includes only 3 atomic species, the lower isotopes of lithium and boron, and the only known isotope of nitrogen. Moreover, these species are all extremely rare. If governed by pure chance this class should contain 25 per cent of all of the atoms in the earth's crust, while it actually contains only about 1/40,000 of this amount. *There are almost no atoms in which the number of nuclear electrons is odd and the number of protons even.*

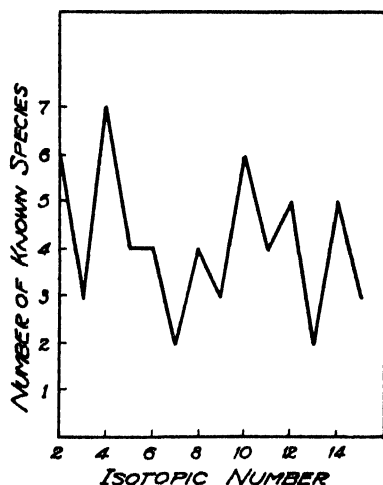


FIG. 20. THE NUMBER OF KNOWN SPECIES OF ATOMS AS A FUNCTION OF THE ISOTOPIC NUMBER

X. FOURTH PERIODIC RELATION: HIGH STABILITY AND ABUNDANCE FOR ATOMS OF ISOTOPIC NUMBER DIVISIBLE BY 4, AND A SECONDARY PERIODICITY OF 2

Among the light elements (atomic numbers 6 to 28) there is a periodicity of 4 in the abundance of the atomic species as related to the isotopic number. This is apparent in figures 18 and 19, and tables 11 and 12.

Furthermore, as is shown plainly by figure 20, there is a periodicity of 2 in the *number of species now known* with respect

TABLE II

Abundance of the atomic species in the meteorites as classified according to their isotopic and class numbers

(Atomic percentage)

ATOMIC NUMBER	SYMBOL	PER CENT	ATOMIC NUMBER	SYMBOL	PER CENT
Isotopic number = 0 Class number = I			Isotopic number = 1 Class number = II		
6	C	0 12	11	Na	0 62
8	O	53 36	13	Al	1 21
12	Mg ²⁴	9 86	15	P	0 06
14	Si ²⁸	14 45	19	K	0 11
16	S ³²	1 42	Total		2 00
20	Ca ⁴⁰	0 97			
Total		80 18			
Class number = IV			Class number = III		
7	N	0 00	12	Mg ²⁵	1 65
			14	Si ²⁹	0 80
			16	S ³³	0 01
			Total		2 46
Isotopic number = 2 Class number = I			Isotopic number = 3 Class number = II		
12	Mg ²⁶	1 65	17	Cl	0 00
14	Si ³⁰	0 20	19	K ⁴¹	0 005
26	Fe ⁴⁴	0 60	Total		0 005
16	S ³⁴	0 03			
28	Ni ⁵⁸	0 31			
Total		2 79			
Isotopic number = 4 Class number = I			Isotopic number = 5 Class number = II		
20	Ca ⁴⁴	0 002	23	V	0 00
22	Ti	0 005	25	Mn	0 06
24	Cr	0 13	27	Co	0 04
26	Fe ⁵⁶	12 10	Total	0.10
28	Ni ⁶⁰	0 19			
Total		12 43			

It is worthy of note that the above percentages make up 99.9 per cent of all of the material of the meteorites, and the similar atomic species, atomic numbers up to 28 and isotopic numbers—1 to 5, make up 99.9 per cent of the earth's surface, which indicates that *no increase in the ratio of electrons to protons (N/P) is able to sufficiently stabilize atoms with a positive nuclear charge greater than 28 as to make such atomic species any considerable factor in the abundance relations. The important feature of this relation is that the abundance does not fall off gradually but suddenly, and remarkably enough the fall in abundance comes just beyond nickel, while nickel is the first of the elements in the region of abundant isotopes. All abundant atomic species have isotopic numbers 4 or less.*

to the isotopic number. The appearance of the plot (14, 23) represented by figure 10 seems to indicate that this periodicity will remain, at least between isotopic numbers 2 and 10, even after much more thorough investigations have been carried out. The periodic relations are summarized in table 13.

TABLE 12

Abundance of the atomic species in the earth's crust as classified according to their isotopic and class numbers

ATOMIC NUMBER	SYMBOL	PER CENT	ATOMIC NUMBER	SYMBOL	PER CENT
Isotopic number = 0 (N/P = 0.5) Class number = I			Isotopic number = 1 Class number = II		
6	C	0.047	3	Li ⁷	0.009
8	O	61.94	5	B ¹¹	
12	Mg ²⁴	1.40	9	F	0.086
14	Si ²⁸	19.40	11	Na	2.64
16	S	0.034	13	Al	6.43
20	Ca	1.88	15	P	0.089
			17	Cl ³⁵	0.025
			19	K	1.42
Total.....		84.70	Total.....		10.70
Class number = IV			Class number = III		
3	Li ⁶	0.0007	12	Mg ²⁶	0.22
5	B ¹⁰		14	Si ²⁹	1.50
7	N		16	S ³²	0.0003
Total..		0.0007	Total....		1.76
Total isotopic number = 1 = 84.70			Total isotopic number = 2 = 12.46		
Isotopic number = 2 Class number = I			Isotopic number = 3 Class number = II		
12	Mg ²⁶	0.21	17	Cl ³⁷	0.007
14	Si ³⁰	0.30	19	K ⁴¹	0.07
16	S ³⁴	0.0006	21	Sc	0.00000
26	(Fe ⁵⁴)	0.09	Total....		0.007
28	(Ni ⁵⁸)	0.006			
Total.....		0.60			

TABLE 12—*Continued*

ATOMIC NUMBER	SYMBOL	PER CENT	ATOMIC NUMBER	SYMBOL	PER CENT	ATOMIC NUMBER	SYMBOL	PER CENT
Isotopic number = 4* Class number = I			Isotopic number = 5 Class number = II			Isotopic number = 6 Class number = II		
10	Ca ⁴⁴	0.05	17	Cl ³⁹		30	(Zn ⁶⁶)	(0.0000×?)
22	Ti	0.28	23	V	0.009			
24	Cr	0.017	25	Mn	0.037	32	(Ge ⁷⁰)	
26	Fe ⁵⁶	1.80	27	Co	0.000			
28	Ni ⁶⁰	0.004	29	Cu	0.0006(?)	36	Kr ⁷⁸	
30	Zn ⁶⁴	0.000						
Total.....		2.15	Total....		0.047			

* It should be noted that in isotopic number 4 the relations of isotopic number 0 are repeated, but with a much smaller abundance. Thus both have a high abundance relative to the adjacent atomic numbers; in both Class I is prominent, and while Class IV has not yet been found in isotopic number 4, it may be present, but only in a very minute quantity. Isotopic number 5 repeats in a lesser degree the relations of isotopic number 1, again with the omission of the secondary class, Class III. While the numbers given in the table may be changed slightly in a few cases by the discovery of new isotopes, the atomic weight and abundance relations indicate that *such changes will not affect the general relations* presented.

TABLE 13
Periodicity of 4 in the isotopic number

ISOTOPIC NUMBER	ATOMIC PERCENTAGE	
	Meteorites	Earth's crust
0	80.1	84.7
1	4.5	12.5
2	2.8	0.6
3	0.005	0.08
4	12.4	2.2
5	0.1	0.05
6	0.000	0.000

XI. THE NEW PERIODIC SYSTEM IN THREE DIMENSIONS (14, 15)

The new periodic system needs at least 3 dimensions in order to exhibit its characteristics. Consider figure 10 to be in a horizontal plane. Then the peaks of figure 18 (or figure 19) may be considered to rise above the respective lines of constant isotopic number (isoneutrons lines). That is if the X- and Y-axes

represent the atomic and isotopic numbers, the Z-axis (vertical may be taken to represent either the abundance or the stability of the atoms. Since there are, however, several types of stability, several such 3 dimensional figures are needed.

XII. FIFTH PERIODIC RELATION: ATOMS OF ODD PROTONIC AND ISOTOPIC NUMBERS (16)

If in figure 10 lines of constant *even* isotopic number are considered, the difference in the atomic number is almost universally 2, and the change of atomic weight 4, between species which are sufficiently abundant to be already discovered.

The number of horizontal steps or layers lying one above the other is from 4 to 7, for atomic numbers higher than 28. *On changing to a consideration of the lines of odd isotopic number and atomic weight, a very different pattern is revealed.* Here it seems common for the change of atomic number along a line of constant isotopic number to be 1, although a change by 2 remains prominent, and instead of from 4 to 7 horizontal steps lying one above another in the region of the heavier elements, there is often only a single step. Also, quite commonly, two steps lie one over the other, but this occurs most generally where there is a transition from a lower to a higher level in going from left to right. In only one case now known do 3 such steps lie above one another, and this is in tin, which is an element of *even atomic number*. Thus far, *more than 2 isotopes have in no case been discovered by Aston or others for any non-radioactive element of odd atomic number.*

While the periodicity of the pattern under discussion is 2 in the isotopic number, it has been mentioned that it is often 1 in atomic number; however, this means a periodicity of 2 in atomic weight. Thus, whenever on a line of constant isotopic number (isoneutronic line) the atomic number changes by 2, which is the more general change for the whole system, the atomic weight changes by 4 and the nucleus adds one α -particle in going from left to right. However, when the change in the atomic number is only 1, as is common when the isotopic number and the atomic weight are odd, the addition is equivalent to a half α -particle, or p_2e . Now, it is of great interest that the formula of any

atomic species of zero isotopic number is $(p_2e)_0$, since zero isotopic number is the fundamental base line of the system, and no nucleus which is complex is represented by a point below this line. Thus the hypothetical group $(p_2e)^+$ seems to play a part which is secondary only to that played by its double $(p_2e_2)^{++}$ which is the α -particle. The writer has suggested the hypothesis that the first step in the formation of an α -particle may be the union of 2 electrons with 1 proton to form the group, which is very difficult to break up, but readily unites with a like group to form an α -particle. Sometimes it may add itself to a complex nucleus. The above relation may mean that it adds on more readily when the atomic weight is odd. However, there are so many other ways in which these nuclei might be formed that little weight should be given to this last hypothesis.

The system followed by the atomic species of odd atomic weight needs expression in a different way from that given above in order that it may be more fully understood. It was supposed by Harkins and Wilson (1) that in general isotopes of elements of even atomic number should have even atomic weights, and elements of odd atomic number should have odd atomic weights. Their supposition has, indeed, been justified. It may be stated, then, that normally elements of odd atomic weight and isotopic number have odd atomic numbers. However, in some exceptional cases the atomic number is even. The question arises, "When does this occur?" This question may be answered empirically by consulting figure 10. This figure shows that in all known instances up to and including atomic number 36, *each atomic species of odd atomic weight and even atomic number lies on the same isoneutronic line (line of constant isotopic number) as, and directly between two adjacent atomic species of next lower and higher odd atomic number.* However the condition for the existence of *very small amounts* of such isotopes *may be that given in the following paragraph for heavier elements.* Thus Be_1^9 , lies adjacent to and between Li_1^7 and B_1^{11} . Here the superscript gives the atomic weight, and the subscript the isotopic number. Se_9^{77} lies in the same way between As_9^{75} and Br_9^{79} ; Kr_{11}^{84} lies between Br_{11}^{81} and Rb_{11}^{85} ; Mg_1^{25} between Na_1^{23} and Al_1^{27} ; and Si_1^{29} between Al_1^{27} and P_1^{31} .

As the number of particles in the nucleus increases, which occurs as the atomic and isotopic numbers rise, it is natural to suppose that the occurrence of abnormal types of atomic species would become more frequent, and the relations of periodicity less stringently followed.⁷ It is, therefore, not surprising that the condition for the existence of species of odd atomic weight, but even isotopic number, should be less exacting when the atomic number rises as high as 50 to 54. Here, in the elements tin and xenon, the condition seems to be that there shall be one adjacent atomic species of the same isotopic number and of odd atomic number. Thus Xe_{21} lies just to the right of I_{21} , but there is no isotope of caesium, of next higher atomic number than xenon, to the right. Also to the right of Xe_{23} is Cs_{23} , but to the left there is no isotope of iodine. While the element to the left of tin (indium) has not been tested for isotopes, the mean atomic weight of indium (which may not be accurate), and the known isotopes of antimony, seem to indicate similar relations for the type of isotopes under discussion, to those found in the case of xenon.

XIII. SIXTH PERIODIC RELATION: PERIODICITY OF TWO IN THE RELATIVE NEGATIVENESS OF NUCLEI WITH RESPECT TO THE ATOMIC NUMBER: PERIODICITY OF TWO IN THE MEAN ISOTOPIC NUMBER

The mean value for the elements of the ratio N/P varies with a general periodicity of 2 as is shown in figures 8 and 9. Here a peculiar relation is found. For the light elements up to atomic number 32 the mean value of N/P (or the isotopic number) is higher in general for elements of odd atomic number, while for heavier elements the higher values occur in general for elements of even atomic number. The extent to which these relations hold is shown by the figures 8 or 9.

⁷ The writer has shown that when the number of particles becomes abnormally large, as in the radioactive species, the regularity again increases, due to the fact that the less stable isotopes here become so unstable as to cease to exist in detectable quantities.

XIV. HYPOTHESIS THAT THE NEGATIVE ELECTRONS IN ATOM NUCLEI ARE ASSOCIATED IN PAIRS

1. *The negative electrons in atom nuclei seem to be usually associated in pairs*

This idea is suggested by the fact that there are two electrons in the α -particle, and that the number is even for almost all nuclei. Furthermore whenever electrons are emitted in the disintegration of the radioactive atoms, there are usually two successive emissions of an electron, and in the remaining cases

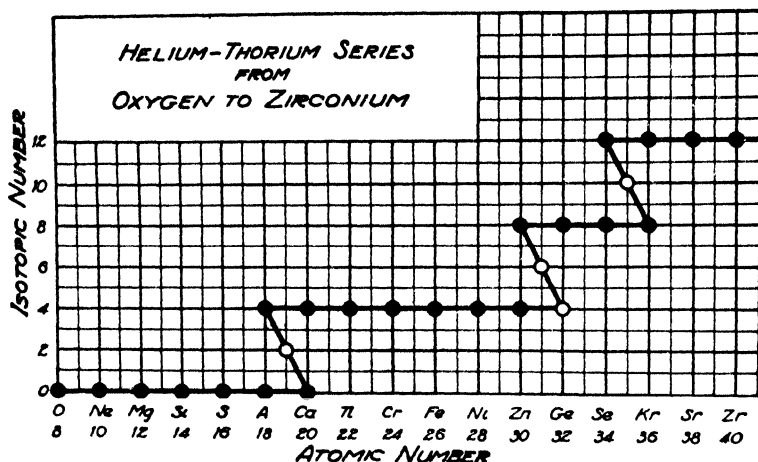


FIG. 21

one electron is emitted just before, and a second just after, the ejection of an α -particle.

About 39 out of 40 known atoms have an even, and 1 in 40 an odd, number of negative electrons. It should be kept in mind, too, that if a nucleus contains 50 pairs of electrons, the presence of one odd electron would give oddness to the total number.

XV. ATOM BUILDING

1. *Are atom nuclei built up from alpha particles?*

a. The data and the periodic relations already presented in this paper give extremely strong support to the idea which led to

the discovery of the relations: that atom nuclei are built up largely from or of alpha particles. Other evidence is outlined below.

b. The radioactive atoms emit alpha particles in most of their disintegrations.

c. A plot of the light atomic species (fig. 21) exhibits exactly the relations of a radioactive series (14), with respect to changes which correspond to the composition of an alpha-particle. Thus in the figure the series of 26 species contains every species but one which could be formed by a change in composition by an alpha-particle. It is remarkable that in this whole series there

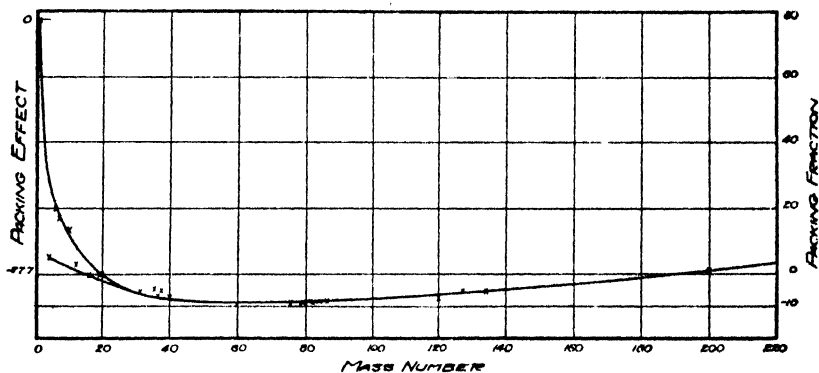


FIG. 22. PACKING EFFECTS IN THE FORMATION OF ATOMS FROM HYDROGEN

is not a single species which does not belong to the helium-thorium series. That is, every one of the 19 species of an isotopic number 0, or which is divisible by 4, but with an odd atomic number, is missing, while every one of the 23 species of even atomic number, except one, has already been discovered. The exceedingly remarkable nature of this relation has not been appreciated.

Three pairs of beta-differences are represented in the figure. It may be noted that each of the intermediate beta-derivatives ($19 K_2$, $31 Ga_6$, and $35 Br_{10}$) is missing, but this is just what is to be expected from the stability relations of the radioactive series, since atoms which give the second of two beta-disintegrations are in general extremely unstable.

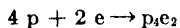
d. A fast-moving alpha particle has been found to attach itself directly to the nucleus of a nitrogen atom.

e. The work of Rutherford (17) and of Rutherford and Chadwick (18) shows that *the light atoms of even atomic number are more stable with respect to a disintegration to give hydrogen than the adjacent atoms of odd number*. This is additional evidence that periodic relation (I) developed by the writer is of fundamental importance.

f. *This difference between odd and even elements is shown also by the packing effects*. It is found that the packing effect for very light atoms figure 22 is larger for even than for odd (Costa (20) and Aston (21)) atomic numbers. This indicates a greater general stability for atoms of even number. It is to be expected that this will be obscured in the heavier atoms, as is the case.

2. *Energy of formation of helium from hydrogen; formation of alpha particles from protons and electrons (helium and hydrogen)*

That the energy which would be liberated if hydrogen were to be converted into helium is very great was shown by Harkins and Wilson in 1915 (19). The amount of energy given off per mol. of helium formed was found to be 2.8×10^{19} ergs or 6.7×10^{11} calories, which is equivalent to 4.57×10^{-6} ergs per α -particle formed. If the reaction were to occur in one step,



as was assumed in the early papers, then the radiation emitted should have a wave length of $4.30 \times 10^{-4}\text{\AA}$, which is almost the value estimated as the wave length of the penetrating or cosmic rays.

However, it seems *possible* that the reaction may occur in two or more steps, since the probability of a sufficiently close approach of 4 protons and two electrons at one instant would be small at any ordinary concentration of protons and electrons.

The intermediate products which seem most probable are: (a) the neutron (pe); (b) a nucleus of isohydrogen (p.e). Thus an alpha-particle could be formed by the meeting of four neutrons and the loss of two electrons; or by the union of two neutrons to

form p_2e_2 , and the loss of one electron to give isohydrogen (p_2e). The alpha particle could then be formed either by the union of two neutrons with an isohydrogen nucleus and the loss of an electron, or by the union of two isohydrogen (p_2e) groups. If neutrons do not form, then it is possible that the primary process is the union of two protons with one electron to form isohydrogen.

It is obvious that if the formation of the helium nucleus occurs in several steps, the wave length of the radiation emitted would be greater than if it were to occur in only one step; and that the wave length emitted may depend upon the kinetic energy of the particles concerned. The union of a proton and an electron at low velocity to form radiation alone would give a wave length equal to $3.3 \times 10^{-8}\text{\AA}$, while their union to form a neutron may be expected to give a wave length of the order of $2 \times 10^{-3}\text{\AA}$.

Thus it seems probable that if the wave lengths of the spectrum of the cosmic radiation could be determined, some basis for a judgment as to the steps which occur would be obtained.⁸

The beautiful work of Paneth and Peters (22) reported recently serves to demonstrate that if helium is formed at all from hydrogen activated by adsorption on a metal catalyst, the rate of reaction is extremely slow, since their extremely delicate tests revealed no helium.

3. Are neutrons used in the building of atom nuclei? (3)

There are two arguments for the idea that neutrons, as well as alpha particles, have been used in the building of atom nuclei.

⁸ Since this was written, Millikan and Cameron (Phys Rev, **31**, 921 (1928)) have found three bands ($\mu = 0.04, 0.08$, and 0.04) in the cosmic radiation. The smallest amount of energy, which according to the Compton-Dirac equation corresponds with one of these coefficients, agrees well with the wave length given above for the formation of helium, but there are also two higher frequencies or lower wave lengths, and the processes which they represent are entirely unknown, although Millikan (Science, **48**, 281 (1928)) seems to imply that such atom nuclei as those of silicon are built up in one direct step from protons and electrons. He also assumes that the nuclei of atoms of iron are built up in the same way. It should be realized that this would mean that 56 protons and 30 electrons, or 86 particles in all meet in space in the minute volume occupied by an atom nucleus or possibly a slightly larger volume, and then unite so suddenly that all of the energy is given off as a single quantum of energy. Other experimental evidence seems to indicate that an atom nucleus has a volume of only about a trillionth that of the atom.

The first of these is that, if neutrons exist, they could approach the positive nucleus of an atom without suffering the great repulsion experienced by an α -particle, or a proton, or any other positively charged particle. This becomes an extremely important factor in the case of the atoms with high nuclear charge. Also, unless it is assumed that the radioactive elements have always existed, it is difficult to see from what source α -particles with sufficient speeds could have come. That neutrons play a part in atom-building was suggested independently by Rutherford (22) and by the writer (3).

The second argument is that the relative abundance of the isotopes of the non-radioactive elements varies only slightly, while the abundance of the elements varies by a factor of the order of a billion. It must be admitted that this argument is partly illusory, since if isotopes vary greatly in abundance the least abundant ones are not discovered, as the sensitivity of the positive ray method is at present probably not better than 1 per cent. Nevertheless, the fact that the known isotopes vary as little in abundance as is the case, is remarkable, especially among the elements of atomic number higher than 29.

If a single neutron is added to an atom nucleus the atomic number remains constant, but the isotopic number increases by unity. Thus magnesium has abundant isotopes of numbers 1 and 2, which may have been formed from the fundamental isotope of number 0, and just the same isotopes are also found in silicon and sulfur. That neutrons have not been discovered, is not a valid argument that they are non-existent, since they would have no known ordinary property except mass. Therefore they could have been discovered only by finding that the chemical analysis of a body does not indicate 100 per cent of known elements. They would not adhere to ordinary substances, and could pass through atoms without noticeable effects except when they approach atom nuclei or electrons with extreme closeness (say to a distance of the order of 10^{-12} cm.). Their concentration in ordinary materials should therefore be very small.

The only evidence which hints at their non-existence is the fact that all known complex nuclei exhibit ratios of N/P between

0.5 and 0.614 whereas this ratio in the neutron would be 1.0, or very much greater than in ordinary stable nuclei. Also this ratio does not exceed 0.572 in the case of any known light atom. This suggests that a single neutron (pe) may unite readily with a proton to form (p_2e) in which the ratio of N/P is 0.5 or that exhibited by the majority of complex nuclei in existence. It may

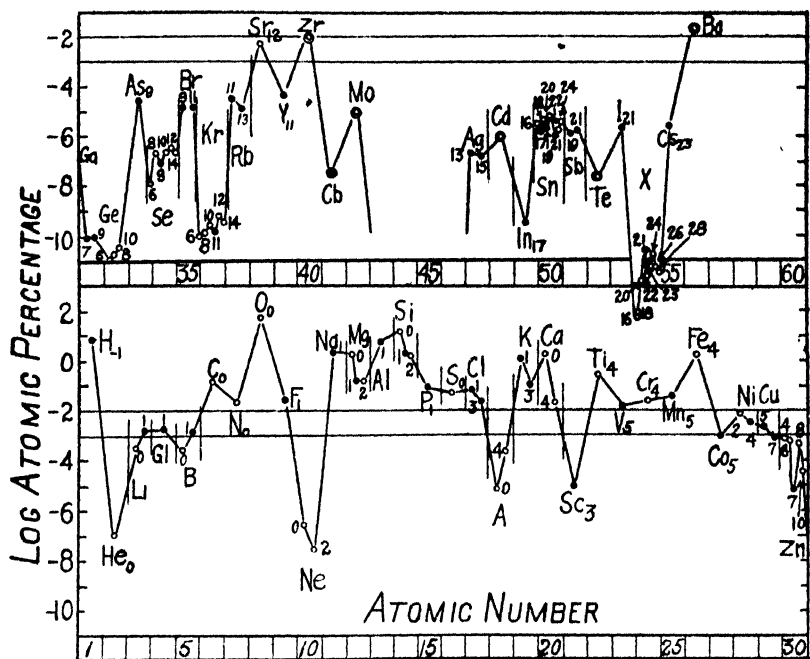


FIG. 23. ABUNDANCE OF THE ATOMIC SPECIES ON EARTH

Where isotopes are given they are designated by their isotopic numbers

also be assumed that neutrons may also unite with complex positive nuclei, but that a large number of neutrons will not exist as such without additional protons in any complex nucleus, but will unite with protons to form more stable groups, such as p_2e , p_3e , etc. It is of interest that the other groups suggested by the composition of complex nuclei are the double neutron (pe)₂, and the group p_2e , which is equivalent to $pe + p_2e$.

4. *Abundance of atomic species on earth*

Figure 23 shows⁹ that there is an enormous difference in the abundance of the various elements in the crust of the earth. The abundance of any element in the figure is the sum of the values given for its isotopes. Obviously the abundance of *known* isotopes of non-radioactive elements cannot vary greatly, since the existence of the isotopes is determined by analyzing a specimen of the element, and the sensitiveness of the methods used thus far is not sufficient to detect small percentages.

XV. FIRST PERIODIC RELATION AND THE BUILDING OF THE EARTH

In closing it may be well to consider the bearing of the periodic system of the atomic species upon other seemingly unrelated theories or ideas connected with the origin of the earth upon which we live.

The familiar saying: "The exception proves the rule," might better be stated: "The exception to one rule may give evidence for another." This seems to be true of the relation that even numbered elements are much more abundant than the adjacent elements of odd number.

The exceptions to this rule give evidence in favor of the theory that the surface of the earth has been built up from small bodies and thus supports the planitesimal hypothesis of Chamberlin rather than the theory of gaseous condensation. Thus it is found that though neon and argon should be extremely abundant if the earth's composition were to represent the general composition of material, they are actually extremely rare. Evidently, they have escaped, which indicates that the earth must have been built from bodies too small to hold these elements by virtue of their gravitational attraction. It is extremely striking that no gaseous element incapable of forming solid or liquid material at ordinary temperatures, is at all abundant on earth. The ordinary gaseous substances, oxygen and nitrogen, seem to have been held mostly as solids or liquids.

Another set of facts agrees with this idea. Both chlorine and

⁹ A similar figure has been used by Aston

nickel are mixtures of two different kinds of atoms. Now it has been found by Baxter for nickel, and by Stone and the writer for chlorine, that the composition of each of these mixtures is exactly the same on earth as it is in the meteorites.

Although chlorine may be separated in the laboratory by very simple physical means, the chlorine on earth does not seem to have changed in its composition in the whole period of existence of the earth's crust.

REFERENCES

- (1) HARKINS AND WILSON, *J. Am. Chem. Soc.*, **37**, 1367 (1915); *Proc. Nat. Acad. Sci.*, **1**, 276 (1915); *Z. Anorg. allgem. Chem.*, **97**, 175 (1916).
- (2) HARKINS, *J. Am. Chem. Soc.*, **39**, 856 (1917).
- (3) HARKINS, *J. Am. Chem. Soc.*, **42**, 1965 (1920).
- (4) HARKINS AND RYAN, *J. Am. Chem. Soc.*, **45**, 2095 (1923).
- (5) BLACKETT, *Proc. Roy. Soc. (London). A*, **107**, 349 (1925).
- (6) HARKINS AND SHADDUCK, *Proc. Nat. Acad. Sci.*, **12**, 707 (1926); *Z. Physik*, **50**, 97-122 (1928).
- (7) HOLUBEK, *Z. Physik*, **42**, 704 (1927).
- (8) DURRANT, *J. Am. Chem. Soc.*, **39**, 621 (1917).
KOSSEL, *Physik. Z.*, **12**, 265-9 (1919).
- (9) GEIGER AND MARSDEN, *Proc. Roy. Soc. (London)*, **82A**, 495 (1909).
GEIGER, *Ibid.*, **83**, 492 (1910); *Phil. Mag.*, **5**, 604 (1913).
RUTHERFORD, *Ibid.*, **21**, 669 (1911).
- (10) FARRINGTON, *Publications* 120 and 151, Field Columbian Museum, Chicago, "Meteorites," 205 (1915); *I. Geol.*, **9**, 623 (1901).
BVISSE, *Mem. Soc. Lits. Sci. Arts l'Aveyron*, **7**, 168.
MEUNER, "Cours. de Geol. Comparee".
SUSS, "The Face of the Earth," (English translation), Vol. 4, p. 543.
- (11) GOLDSCHMIDT AND THOMASSEN, *Videnskaps Skrift. I. Mat.-naturw. Kl.*, 1924, No. 5.
- (12) HARKINS, *Phys. Rev.*, **15**, 79-94 (1920).
- (13) HALL, N. F., *J. Am. Chem. Soc.*, **39**, 1616 (1917).
- (14) HARKINS, *Phil. Mag.*, **42**, 306 (1921).
- (15) HARKINS, *J. Am. Chem. Soc.*, **43**, 1038 (1921).
- (16) HARKINS, *Ibid.*, **45**, 1426 (1923).
- (17) RUTHERFORD, *Phil. Mag.*, **37**, 538 (1919).
- (18) RUTHERFORD AND CHADWICK, *Phil. Mag.*, **42**, 809 (1921), **44**, 417 (1922).
- (19) HARKINS AND WILSON, *Phil. Mag.*, **30**, 723 (1915).
- (20) COSTA, *Ann. Phys.* (10), **4**, 425 (1925).
- (21) ASTON, *Proc. Roy. Soc. London, A*, **115**, 487 (1927).
- (22) PANETH AND PETERS, *Ber.*, **59**, 2039 (1926).
- (23) RUTHERFORD, *Proc. Roy. Soc. London*, **42**, 1964 (1920).

ATOMIC STRUCTURE AS REVEALED BY CRYSTAL REFLECTION OF X-RAYS

G. E. M. JAUNCEY

Department of Physics, Washington University, St. Louis, Missouri

In 1912 Laue discovered that crystals act towards x-rays as a three dimensional diffraction grating. W. L. Bragg simplified the mathematics of Laue's explanation of his discovery by showing that the Laue spots could be explained by reflection from the various planes in a crystal in which the atoms may be considered as being arranged. Referring to figure 1A, the white circles represent sodium atoms and the black circles chlorine atoms in a crystal of rocksalt. It is easily seen that different planes of atoms can be chosen in the crystal as represented in figure 1B. It is seen that the (100) and (110) planes in the crystal contain both chlorine and sodium atoms in each plane. Such planes may be called even planes. On the other hand there are planes like the (111) planes which contain either all chlorine or all sodium atoms in a plane—the chlorine and sodium planes alternating with each other. Such planes may be called odd planes.

Referring again to figure 1, we see that the grating space d varies according to the sets of planes considered. The Bragg law of reflection is

$$n\lambda = 2d \sin \theta \quad (1)$$

By cutting a crystal so that its face is parallel to the (100) planes and determining the angle θ for first order reflection of a known wave length, the value of d_{100} , the grating space for the (100) planes can be found. Likewise, by cutting the crystal so that its face is parallel to the (110) planes and determining the angle θ for first reflection of the same wave length, d_{110} can be found. Similarly, by cutting the crystal so that its face is parallel to the

(111) planes, d_{111} can be found. From the values of d_{100} , d_{110} , d_{111} , the crystal structure, which means the arrangement of the atoms in the crystal, may be determined. Debye and Scherrer and also Hull have devised what is known as the powdered crystal method. In this method we obtain in one and the same exposure and on one and the same photographic film a

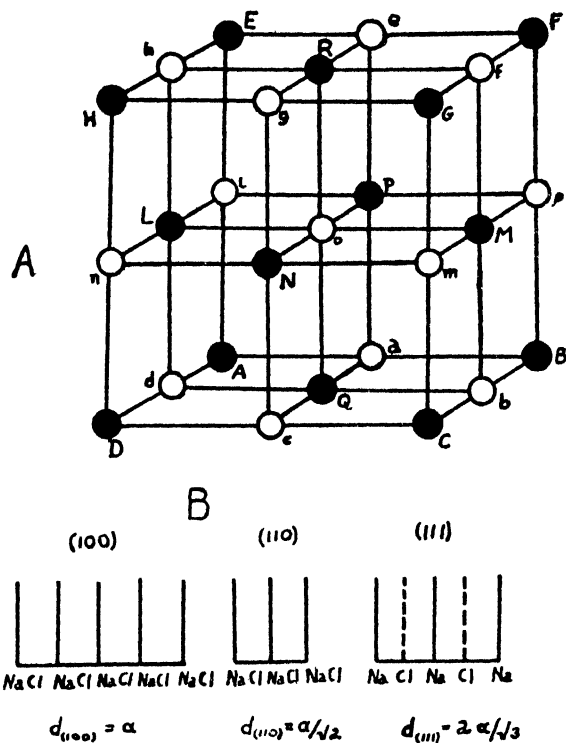


FIG 1

record of the reflections of a given wave length from all the sets of planes. From the arrangement of the lines showing the reflections it is possible to obtain the crystal structure.

Our interest in the present paper, however, is not so much in the arrangement of the reflections as in the intensities of the various reflections. From the arrangement of the reflections the crystal

structure is found, while from the intensities we are able to get at the atomic structure, or the arrangement of the electrons in the atom. In 1921 W. L. Bragg and his pupils began a study of the intensity of the reflected x-rays in various orders and from various planes in rocksalt. The arrangement is shown in figure 2. A spectrum line of a certain wave length is isolated by the first crystal and reflected by the second crystal of rocksalt. Since rocksalt is not a perfect crystal but consists of a mosaic of small crystals whose axes make small angles with each other and since each one of the small crystals must be brought into the position to give reflection according to the Bragg law, the method used by Bragg is to rotate the second crystal slowly through the position

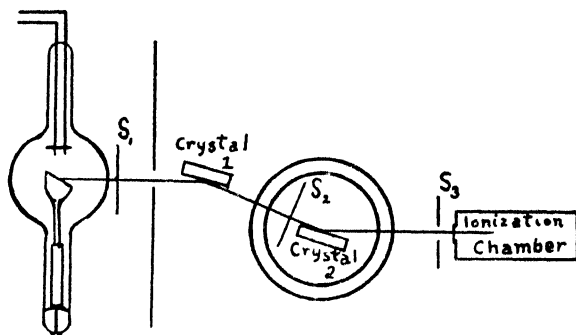


FIG. 2

of reflection with the angular velocity ω . The reflected rays enter an ionization chamber and the ionization current is registered by the electrometer. If E is the total deflection of the electrometer produced while the second crystal is rotating through the position of reflection and if I is the intensity of the rays incident on the second crystal determined by removing the second crystal and turning the ionization chamber into such a position as to receive the incident rays from the first crystal and obtaining the deflection of the electrometer in unit time, then $E \omega / I$ is known as the integrated intensity.

According to C. G. Darwin, the integrated intensity is given by

$$\frac{E \omega}{I} = \frac{N^2 e^4 F^2 \lambda^1}{2m^2 c^4 \sin 2\theta} \cdot \frac{(1 + \cos^2 2\theta)}{2 \mu} \quad (2)$$

where N is the number of molecules of sodium chloride per cubic centimeter if rocksalt is used, θ is the grazing angle of incidence at which a reflection takes place according to Bragg's law, λ is the wave length of the x-rays, e and m are the charge and mass of the electron respectively, c the velocity of light (inserted because e is in electrostatic units), μ is the linear absorption coefficient of the x-rays in the crystal and F is a measure of the diffracting power of the atom. All the quantities in equation (2) can be measured or are known except F so that equation (2) can be used to determine F . This quantity F is now known as the atomic structure factor. It should be stated here that μ is not the ordinary linear coefficient of absorption, because W. H. Bragg has shown that when a crystal is in a position to reflect x-rays its absorption coefficient is abnormally high. The increase in the ordinary absorption coefficient is termed by W. L. Bragg the extinction coefficient. Methods for the determination of the extinction coefficient are described in the papers of W. L. Bragg and also in Chapter V of A. H. Compton's "X-rays and Electrons" and will not be further discussed here. Returning to the consideration of F , it is found that when the experimental values of the integrated intensity are substituted in equation (2), and the values of F calculated, F is a function of the order of reflection and of the set of planes—(100), (110), (111) and so on—for which the integrated reflection is measured. However, since $\sin \theta$ is determined by the order of reflection and by the set of planes, F is usually given as a function of $\sin \theta$. Referring to figure 3, we see that the F values which are plotted against $\sin \theta$ fall upon two smooth curves. Curve I is obtained from the integrated reflection for the even planes, such as the (100) and (110) planes, while curve II is obtained from the integrated reflections of odd order from the odd planes such as the (111) planes. The F values obtained from the integrated reflections of even order from the odd planes fall upon the upper curve. This is because each of the even planes consists of sodium and chlorine atoms and these two types of atoms coöperate in the reflection process. Hence the F curve obtained in this case is called the (Cl + Na) curve. The odd planes, however, consist either of all chlorine or

all sodium atoms. The grating space for such planes is not the distance between one plane and the next plane but the distance from one chlorine to the next chlorine plane or the distance from one sodium to the next sodium plane. In the first order reflection, therefore, the rays reflected from the second chlorine plane

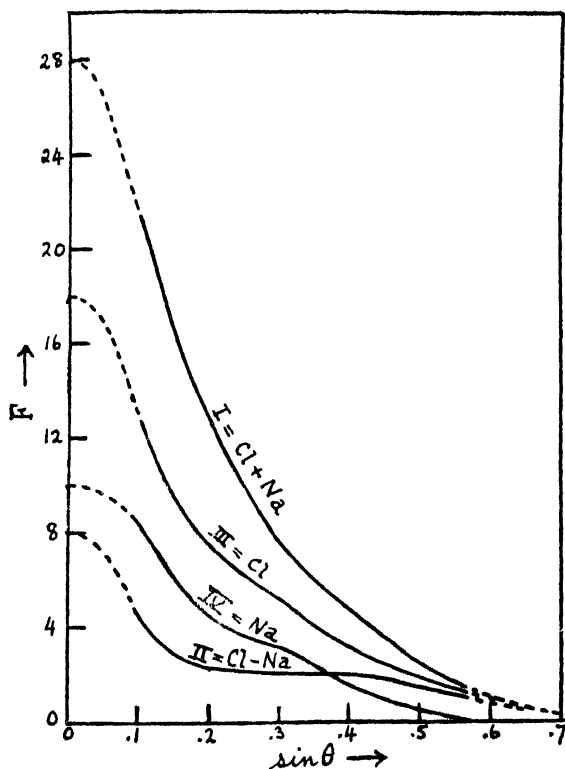


FIG. 3

are one wave length behind the rays reflected from the first chlorine plane, but in between the two chlorine planes is a sodium plane and the rays reflected from this plane are one half wave length behind the first chlorine plane and, therefore, are exactly out of phase with the rays reflected from the chlorine planes. If the reflection from a sodium plane were equally as intense as from

a chlorine plane there would be no reflection in the first order. However, reflection from the sodium planes is not as strong as that from the chlorine planes and what is observed is the difference of the intensities from the chlorine and sodium planes. The F curve obtained for this case is, therefore, called the $(\text{Cl} - \text{Na})$ curve. In the case of even order reflection from the odd planes the reflected rays from the sodium planes are in phase with the reflected rays from the chlorine planes and the F values for this case, therefore, fall on the $(\text{Cl} + \text{Na})$ curve. Obviously, if we have the F values for $(\text{Cl} + \text{Na})$ and for $(\text{Cl} - \text{Na})$, and since

$$\text{Cl} = (1/2) \{(\text{Cl} + \text{Na}) + (\text{Cl} - \text{Na})\} \text{ and } \text{Na} = (1/2) \{(\text{Cl} + \text{Na}) - (\text{Cl} - \text{Na})\} \quad (3)$$

we can obtain the F curves for Cl and Na alone. These are shown as curves III and IV respectively in figure 3. The full portions of the curves in figure 3 are obtained from experimental values of the integrated intensity, while the broken portions are extrapolations.

We shall now consider the meaning of the F curves. If the diffracting centers in the crystal were exactly in the geometric crystal planes, the F values would be constant with the angle and the F curves would be parallel to the $\sin \theta$ axis. The reason why crystals reflect x-rays and amorphous substances do not reflect x-rays is that the diffracting centers in a crystal are arranged in planes which are regularly spaced from one another while in an amorphous substance the diffracting centers are arranged at random and there are no definite planes. Any factor which tends to make the planes of diffracting centers in a crystal indefinite or diffuse will reduce the integrated intensity of the reflected x-rays and hence by Darwin's formula will make F smaller. The fact that the F curves in figure 3 are not parallel to the $\sin \theta$ axis, therefore, means that the planes are to some extent diffuse. What factors operate to cause diffuseness of the planes? Referring to figure 4, we see that if the atoms are vibrating due to their temperature then these thermal vibrations will cause the atoms to make random excursions from the geometric planes of the crystal. Imagine an instantaneous photograph taken of the crystal and the picture will be like that shown

in figure 4. The atoms will be distributed about the crystal planes as shown, some above and some below the planes. As the temperature increases the diffuseness of the planes will increase and the intensity of reflection will decrease because the crystal will become more and more like an amorphous substance. Debye has discussed this effect of the temperature on the intensity of reflection and has obtained a factor $\exp(-b \sin^2 \theta)$ for the temperature effect, where b is a function of the temperature, the specific heat of the crystal and the wave length but is a constant for a given temperature and wave length. The effect of an atom—or, rather, an electron, since it is the electrons which are

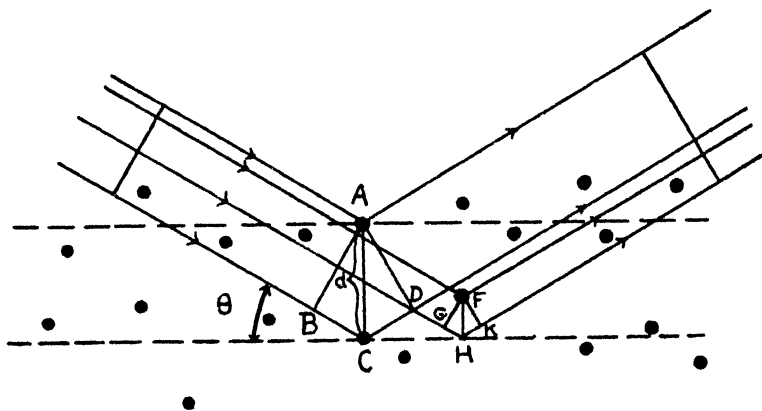


FIG. 4

the diffracting centers—being displaced from a plane is that the rays diffracted by such an electron are slightly out of phase with the rays diffracted by an electron in the plane and hence we have partial interference and a reduction of the intensity of the reflected rays and therefore a reduction of the corresponding F value. Consider the electron at F in figure 4. This is displaced a distance $FH = z$ from the geometric crystal plane. The difference of path between a ray diffracted by this electron and another ray diffracted by an electron in the plane is $2z \sin \theta$. The difference of path is therefore $(2z \sin \theta)/\lambda$ wave lengths. Since each wavelength difference of path introduces a phase difference of 2π ,

we see that the phase difference between a ray diffracted by the electron at F and another ray diffracted by an electron in the plane is $2\pi(2z \sin \theta) / \lambda$ or $(4\pi z \sin \theta) / \lambda$. Hence, the displaced electron is equivalent as far as the diffraction effect is concerned to a fraction $\cos \{(4\pi z \sin \theta) / \lambda\}$ of an electron in the plane. If we add the values $\cos \{(4\pi z \sin \theta) / \lambda\}$ for each electron in a cubic centimeter of the crystal and divide by the number of the molecules in a cubic centimeter, we obtain the effective number of electrons in a molecule of NaCl. Thus, in figure 3 we see that F for (Cl + Na) at $\sin \theta = 0.1$ is 21.4. This means that at this angle the molecule of NaCl which really has 28 electrons acts

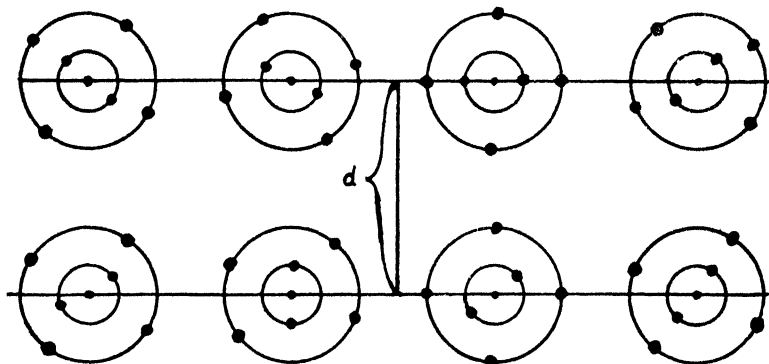


FIG. 5

as if it had only 21.4 electrons as far as reflection of x-rays at this angle is concerned. This is because the 28 electrons are not all in a crystal plane but are displaced by various amounts from the plane and due to partial interference among themselves act as if they were 21.4 electrons in a crystal plane. Similarly, the 17 or 18 electrons in chlorine act as though they were only 12.8 electrons in a crystal plane at an angle of reflection given by $\sin \theta = 0.1$.

Another factor which causes diffuseness of a crystal plane is the fact that the electrons of an atom are themselves distributed in the atom at various distances from the nucleus. It should be remarked here that the nucleus takes no observable part in

the scattering of x-rays. The effect of distribution of the electrons in the atoms is shown in figure 5, in which is shown a fictitious crystal made up of carbon atoms with 2 *K* electrons and 4 *L* electrons revolving in circular orbits. These orbits are not only in the plane of the diagram as shown in figure 5, but also may make any angle with the plane of the diagram. Obviously, this again has the effect of making the crystal planes diffuse. The diffuseness is, therefore, caused by the Debye temperature effect and by the structure of the atom. The atomic structure factor curves, or *F* curves as they are called in the literature, give us an insight into these two effects. When both effects are operative, we have $F = f \exp(-b \sin^2 \theta)$ where *f* is a factor due to the structure of the atom. By comparing the *F* values at a low temperature such as that of liquid air with the *F* values at the same angles at a higher temperature, it is possible to obtain the value of *b* at the higher temperature. For reflection of molybdenum *K*_α x-rays ($\lambda = 0.71$ angstroms) from rocksalt at 20°C. the value of *b* is 2.39. Hence, the true atomic structure factors or *f* values can be disentangled from the temperature factor.

A. H. Compton has shown that the true atomic structure factor or *f* value for an electron moving in a circular orbit of radius *a* is

$$\frac{\sin \{(4\pi a \sin \theta)/\lambda\}}{(4\pi a \sin \theta)/\lambda}$$

If this is summed for each electron in the atom, we obtain the *f* value for the atom. Or working backwards, if we measure an *F* value then after correcting for the temperature factor we can use the formula to obtain *a*, the radius of the orbit. Of course, this is for an ideal case of an atom with only *K* electrons. In the actual case, the electrons are not all in one orbit, and the problem is more complicated. For electrons in two different orbits such as *K* and *L* circular orbits, assuming that we know the numbers of electrons in each orbit, we could by finding the *F* values for two values of $\sin \theta$ and thence obtaining the *f* values calculate the radii of the *K* and *L* orbits. However, the actual problem is even more complicated than this, because the orbits are not all

circular but some orbits are elliptical and perhaps even more complicated than elliptical. Also, we do not know the numbers of electrons in each group of electrons. However, if we take enough F values at enough different angles, we are able to solve the problem. Duane and later A. H. Compton have given us a general method for analyzing the F values and obtaining therefrom the structure of the atom. This method is known as the Fourier analysis method because it makes use of the same mathematical

TABLE I
F values in the electron distributions of Cl and Na
 $D = 3.25$ ångströms
 $\sin \theta$ for first order = 0.109

n	F_n		n	F_n	
	Cl	Na		Cl	Na
1	13.60	8.80	5	2.10	0.10
2	7.84	5.46	6	1.15	0.00
3	5.15	3.05	7	0.31	
4	3.40	0.72			

principle as used in analyzing any periodic motion into its constituent simple harmonic motions. The formula for it is as follows:

$$U = (8\pi r/D^2) \sum_{n=1}^{n=\infty} n F_n \sin (2\pi nr/D) \quad (4)$$

where Udr is the number of electrons between two spheres of radii r and $r + dr$, the centers of the spheres being at the center of the atom, F_n is the atomic structure factor for the n th order reflection from a set of planes whose grating space is D , and r is the distance from the center of the atom. U is called the radial distribution of the electrons in the atom and is usually given in electrons per ångström. It will be noticed that the right side of equation (4) is a summation of a series. It is a very cumbersome formula and the labor of applying it is such that a calculating machine is necessary. In making the calculations for U it is often more convenient to use values of F_n read off from the F

curves of figure 3 rather than the experimental F values. This is justifiable so long as the F values are read off at equal increments of $\sin \theta$, for instance, at $\sin \theta = 0.1, 0.2, 0.3$, and so on, respectively. However, when this is done, a fictitious value of the grating space D must be used. The F values used in the calculation of the electron distributions in chlorine and sodium are shown in table 1.

These are values obtained by Havighurst when molybdenum K_α x-rays are reflected from rocksalt. From these values Havighurst obtains the electron distribution, or U curves shown in

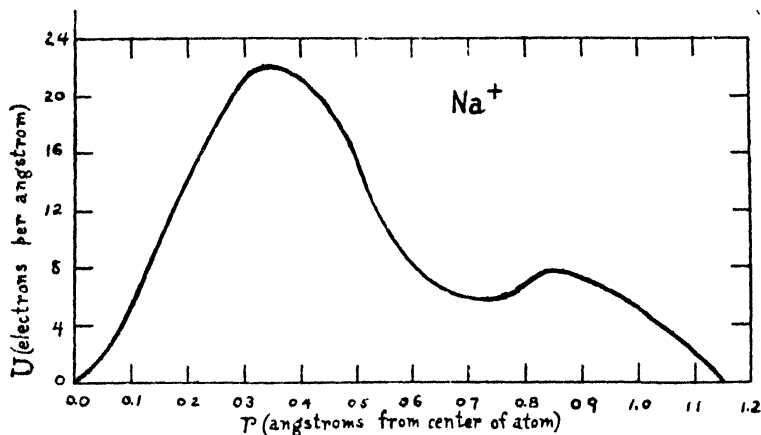


FIG 6

figures 6 and 7. In table 1, the values of F_6 and F_7 for chlorine and of F_6 for sodium are extrapolated. This fact, therefore, introduces some uncertainty into the U curves shown in figures 6 and 7. The position of the maximum of each of the humps gives the distance of each different group of electrons from the center of the atom, while the area under each of the humps gives the numbers of the electrons in the respective groups. Thus, in the sodium atom there are 8 $K + L$ electrons at 0.33 angstrom and 2 L electrons at 0.85 angstrom from the center of the atom. In the chlorine atom, there are 10 $K + L$ electrons at 0.27 angstrom, 5 M_{III} electrons at 0.75 angstrom and 3 M_I electrons at

1.22 angstroms from the center of the atom. The *K* and *L* electrons are merged together because of the heat vibrations. If *D* in equation (4) is made indefinitely great and the increments of $\sin \theta$ indefinitely small, the right side of equation (4) takes the form of a Fourier integral. Applying the Fourier integral method of analysis to the temperature factor alone, Mr. Claus of Washington University and the writer have found that the most probable displacement of the heat vibrations is 0.175 angstrom. Since the radii of the *K* orbits for both sodium and chlorine are thought to be much less than 0.175 angstrom, the effect of the thermal vibrations is to obliterate any evidence for a separate *K* orbit in these analyses. The thermal vibrations have the

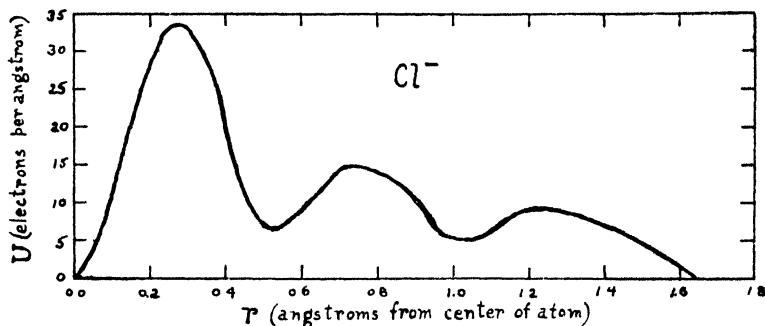


FIG 7

further effect of broadening the humps due to the *L* and *M* electrons, and of shifting the maxima somewhat further out from the center than is the true distance of the *L* and *M* electrons from the center of the atom. It should be remarked here that a *U* curve gives the electron distribution relative to the average center of the atom, the true center of the atom vibrating about the average center due to the temperature effect and carrying the electrons with it.

The exact interpretation of the *F* curves is still in the process of being worked out. It seems to the writer that the Compton effect must somehow be taken into account. In fact, the *F* curves as given by Havighurst have a shape which cannot be

obtained from the classical theory while the whole of this present paper is based on the classical theory. The situation at the time of writing this paper is that more work is needed on the experimental side in order to obtain more accurate F values and also on the theoretical side to show definitely what effect, if any, the Compton effect, which has to do with the change of wave length on the scattering of x-rays, has on the Fourier analysis method and the resulting electron distribution curves. Nevertheless, the writer believes that the method of unravelling atomic structure as outlined in this paper has great promise of being one of the most useful and direct means of adding to our knowledge of the inside of the atom.

REFERENCES

- P. DEBYE, *Ann. der. Physik*, **43**, 49 (1914)
C. G. DARWIN, *Phil Mag.*, **27**, 315 and 675 (1914)
A. H. COMPTON, *Phys. Rev.*, **9**, 29 (1917)
BRAGG, JAMES AND BOSANQUET, *Phil Mag.*, **41**, 309 and **42**, 1 (1921).
C. G. DARWIN, *Phil. Mag.*, **43**, 827 (1922).
W. DUANE, *Proc Nat Acad. Sci.*, **11**, 502 (1925).
J. A. BEARDEN, *Phys. Rev*, **27**, 796 (1926).
BRAGG, DARWIN AND JAMES, *Phil Mag*, **1**, 897 (1926).
R J HAVIGHURST, *Phys. Rev*, **28**, 869 and 882 (1926); **29**, 1 (1927)
G. E. M. JAUNCEY, *Phys. Rev*, **29**, 757 (1927)
JAMES AND FIRTH, *Proc Roy Soc., A*, **117**, 62 (1927)
WALLER AND JAMES, *Proc Roy. Soc , A*, **117**, 214 (1927).
R J HAVIGHURST, *Phys. Rev.*, **31**, 16 (1928)
JAUNCEY AND CLAUS, *Phys. Rev*, **31**, 717 (1928).
JAUNCEY AND CLAUS, *Phys. Rev*, **32**, 12 (1928).
A. H. COMPTON, *X-Rays and Electrons*, Chapter V.

DISSOCIATION OF MOLECULES AS DISCLOSED BY BAND-SPECTRA

KARL K DARROW

Bell Telephone Laboratories, New York City

The evidence which is furnished by band-spectra in regard to the dissociation of molecules is in essence like that which line-spectra provide concerning the ionization of atoms, for the two phenomena are very similar—so much alike, indeed, that it is suitable for convenience of language as of thought to refer to ionization as a special form of dissociation. In both cases there is, in the limit, separation of the individual particles of a gas—molecules or atoms, as the case may be—into pairs or larger sets of disconnected particles. In both cases there are intermediate steps between the normal state of the undisturbed molecule or atom, and the state of its complete resolution into fragments altogether free from one another.

Total separation may often be detected by physical or chemical means, as for instance when the presence of helium ions in helium gas is shown by electrical experiments, or the presence of hydrogen atoms in molecular hydrogen gas is proved by one or another of the well-known chemical tests for atomic hydrogen. Molecules or atoms in one or another of the intermediate, the so-called *excited*, states may also be detected occasionally by physical or chemical means, as when excited mercury atoms are disclosed by mixing the mercury vapor with hydrogen gas and applying a test for atomic hydrogen, which excited but not normal atoms produce by impact against hydrogen molecules. Many of the phenomena of catalysis, indeed, are probably indications of the presence of excited atoms or molecules.

Line-spectra however furnished the earliest comprehended evidence for the existence of excited states; they still provide the most accurate way of locating these states, and locating the state

of ionization. Band-spectra likewise give by far the most accurate way of discovering and locating the intermediate states of a molecule which, so to speak, are steps between its normal unperturbed condition and its complete disaggregation into atoms. It may be superfluous to remark on the importance of such measurements. Even if one were interested only in the attainment of complete dissociation, one could not fail to be instructed by the study of the steps which lead to it; and these steps themselves acquire extreme significance from the fact that in all probability they correspond to various degrees of what is called *activation*. As for complete dissociation, it is discernible from band-spectra as complete ionization is from line-spectra; and here also the accuracy of the spectroscopic method surpasses by far that of any other.

To speak of *locating* excited states or the state of dissociation, or of *accuracy* in measuring them, may give at first a misleading impression. It is not primarily a question of measuring minute quantities of the products of dissociation; rather it is a question of determining the value of energy which on being communicated to the molecule or atom is just sufficient to break it into fragments. The spectra reveal that there is such a critical value of energy, a fact which previously may have seemed reasonable enough but could scarcely have been proved. It is a great advantage of the spectroscopic method, indeed, that the fact of ionization or dissociation, the evidence that it requires a definite minimum amount of energy, and the value of this amount of energy, are all simultaneously furnished by the single spectrum photograph. Much the same thing may be said of the location of excited states. Each is characterized by a definite excess of energy over the energy-value of the molecule or atom in its normal state. To demonstrate an excited state by analysis of a band-spectrum is the last step but one before determining its energy, unless indeed it be the very last step; a complete analysis of any spectrum involves the determination of the energy-values, or, as I shall say, the locations, of all of the excited states and of all of the states of dissociation of the particle which is the carrier of that spectrum. Or, to put the same idea in words slightly

different: complete understanding of any spectrum would entail a knowledge of all of the modes of dissociation possible to its carrier, and of all the intermediate steps leading towards these various modes.

Complete knowledge in any field is, however, rare; and it is fortunate that in line-spectra as well as in band-spectra, the lines or the bands—as the case may be—are sometimes so arranged that a particular series of excited states, culminating in a particular mode of dissociation, is almost obvious to even the uninstructed onlooker. The absorption spectrum of cool sodium vapor, for instance, consists of a striking sequence of dark lines. Successive members of this sequence approach more and more closely to one another as one follows the procession from the yellow toward the ultraviolet, until it is evident that they are “converging” upon a limit-frequency which can be estimated with modest accuracy by visual inspection, although the gradual decrease in the distinctness of the lines as they draw towards the convergence point precludes a definite sighting of this latter. The fact that these lines appear in the absorption spectrum of cool unstimulated vapor—that is to say, when the vast majority of the sodium atoms are necessarily in their normal state—shows that each stands for the transition from the normal into some one of the excited states of the sodium atom. The excited states, or more precisely their energy-values, form a converging sequence as the lines form a converging sequence. If I denote by $\nu_1, \nu_2, \nu_3, \dots$ the frequencies of the various lines in order of increasing magnitude, and by ν_∞ their convergence-frequency; then the energy-values of the corresponding excited states, relatively to the energy of the normal state taken as zero for reference, are equal to $h\nu_1, h\nu_2, h\nu_3 \dots$; and the energy corresponding to the limit of the sequence of excited states, referred to the same zero, is equal to $h\nu_\infty$.

Now this limit of the sequence of excited states is the state of ionization; and the energy-value $h\nu_\infty$ is the energy required to effect ionization. It is the so-called *ionizing potential*; but some such term as *heat of ionization* or *work of ionization* is preferable, for the sake of the analogy with dissociation of molecules. For,

be it recalled, ionization is that special type of dissociation in which the fragments are a free electron and an atom (or, it might be, a molecule) devoid of one of its normal quota of electrons.

Such a definition evidently contains an element of ambiguity. Though a free electron is presumably a definite object—at least, we have not yet found it necessary to postulate two or more varieties of free electrons—there may well be several kinds of ions. It is conceivable that the sodium atom, deprived of a single electron of its normal quota, might assume any one of a great number of different configurations, each corresponding to a different energy-value, and therefore requiring a different value of work of ionization. In other words there might be many distinct ways of ionizing a sodium atom, each involving the expenditure of a distinct amount of energy, and each resulting in a distinct type of sodium ion.

Such is not only conceivably, but actually, the case. We have only to recall the various types of ionization which give rise to the lines of x-ray spectra, and which are described as extraction of a K-electron, extraction of a L_I electron, and so forth. These are comparatively easy to distinguish from what is commonly called ionization without a qualifying word—detachment of the “loosest” or “valence” electron; for they require much more energy. On passing to the “two-valence-electron” atoms such as calcium, however, we find not only these modes of ionization which involve extraction of inner electrons and high values of ionizing-potential, but also others which lie very close to the minimum ionizing-potentials. Or, to express the facts better, we may say that there is not one, but a group, of low ionizing-potentials. In the optical spectrum of such an element, there are sequences of lines disclosing sequences of excited states which converge upon two distinct energy values, corresponding to two distinct amounts of work of ionization, corresponding to two different states of the positive ion which is one of the two fragments into which the atom is resolved. With elements still farther to the right in the Periodic Table, the multitude of possible and not-very-different ionizing-potentials may be great. This I mention in order to prepare for a fact which is important in the analysis of band-spec-

tra; to wit, that the state of the atoms into which a molecule is resolved by dissociation need not always be the same—that it may be different when the dissociating agency is different—that in particular it need not by any means be the normal state of the atoms in question. Precisely this variability of the fragments resulting from dissociation is one of the most interesting features of the entire field.

Now in the most easily interpreted of the band-spectra—such for instance as that of iodine—a not-too-excellent spectroscope discloses a converging sequence¹ of broad hazy lines. The reason for specifying that the instrument be not too good is simply that with a high resolving power each one of these broad hazy lines—which in fact are *bands*—is discovered to be a rich group of fine lines. Likewise in the case of sodium a good spectroscope reveals that each “line” of the principal sequence or series is in truth a pair of lines. In both cases this “fine-structure” is of the greatest interest and importance; in the former case it reveals the rotations of the molecules. In both cases, however, it is irrelevant to our present interest, at all events for a first examination of our topic; and it is better to ignore the fine-structure of the molecular bands as we have ignored that of the atomic lines, considering each band as a unit as we have considered each sodium line as a unit—that is to say, as the sign of a transition between the normal and an excited state of the molecule. Then the converging series of bands in the absorption spectrum of (say) iodine reveals a converging sequence of excited states of the iodine molecule; and the limit ν_{∞} on which they converge, multiplied by h , is the energy-value of the state of dissociation to which this sequence of states leads up—the work of dissociation, of this particular mode of dissociation, reckoned from the normal or initial state of the molecule.

Before entering further into detail concerning these converging sequences, we should consider another sign of dissociation: the presence of a continuous region of absorption extending over a

¹ I continue to use the word “sequence” for the sake of the analogy with the previous case; but in the technical language of the analysts of band-spectra it has a different meaning.

relatively wide range of the spectrum. In the ideal case—such a case as is approached with sodium—the region of continuous absorption should commence at the convergence frequency of the sequence of absorption lines, and extend thence away from the lines toward higher frequencies. Such a phenomenon in an atomic spectrum would signify that light of any frequency—say ν_* —within this region, is able to extract an electron from the atom, conferring upon it—more properly, upon the system of ion and free electron—kinetic energy of translatory motion in amount $(h\nu_* - h\nu_\infty)$. The region of absorption would be continuous, because there is nothing in quantum-mechanics—so far as we know—to limit kinetic energy of unperiodic translatory motion to discrete “permitted” values.

Now there is a continuous zone of absorption in the spectra of sodium and potassium, lying in the main beyond the convergence-frequency of the principal series; and if we are willing to leave for future explanation the fact that this zone intrudes upon the line-series instead of stopping sharply at its limit, we may assume that it is indeed the sign of ionization with transfer of extra kinetic energy to the separated particles. Likewise, there are continuous zones of absorption in the band-spectra of many molecules, some of which lie around and beyond the convergence-limit of a sequence of bands, while others stretch towards long waves as far as the boundary set by the limitations of the apparatus, so that it cannot be told whether or not they adjoin band-sequences. If we assume that each of these continuous regions of absorption is a sign of dissociation with transfer of extra energy to the separated particles, then in these we have an additional source of evidence for dissociation and even of estimates of the energy required to dissociate.

Suppose then that we have collected and tabulated² a number of values of work of dissociation computed either from the short-wave limits of sequences of bands or sequences of excited states, or from the long-wave limits of continuous regions of absorption; can we obtain some independent confirmation of these values?

² Compare such a tabulation in the article of H. Spöner (*Ergebnisse exakt. Naturwiss.* 6, 75–103 (1927))

The obvious step is to compare them with the values of work of dissociation—under these circumstances more often called *heat of dissociation*—obtained by conventional thermal methods. We then discover, first, that the best of the values deduced from the analysis of spectra—taking it for granted, of course, that the work of dissociation is truly the product of h into the convergence-frequency of a sequence—are determined with much greater accuracy than the thermal methods are capable of giving.³ To take one instance only, but the most important; there is no other estimate of the heat of dissociation of molecular hydrogen for which anywhere nearly so small a numerical uncertainty is claimed as for the value—good at least to two per cent—deduced from analysis of the band-spectrum of hydrogen gas. We cannot therefore compare pairs of values of presumptively equal exactness; there is a relatively wide range about any accepted “direct” value, anywhere within which the “spectroscopic” value may be without compelling us to concede a disagreement.

Nevertheless there are cases in which actual undeniable disagreement occurs; and it is most important to observe, that out of these disagreements we infer the principle already stated—that the mode of dissociation effected by light need not immediately result in the formation of two *normal* atoms. What the thermal methods yield, in fact, is the net amount of energy expended upon the gas in the course of transforming its molecules completely into pairs (for simplicity I am speaking as though all molecules were diatomic) of normal neutral atoms. This may mean that in the apparatus used in the thermal experiments molecules are always dissociated in that way. But it need not, and probably does not, mean anything of the sort. If the molecules are first dissociated into excited atoms, and subsequently these atoms pass over into their normal state, then the energy which in so doing the atoms release is restored to the environment; and what is measured is the difference between the energy originally spent in the dissociations, and this fraction which is returned. The spectroscopic method, however, yields the amount of energy

³ Special methods for locating the convergence frequency must of course be employed.

which was spent in effecting that primary dissociation, with no deduction on account of the fraction thereof which may be restored by the resulting atoms in the course of their future evolution.

Now if in any particular case the state of affairs is actually such as I have just described, we may expect the "spectroscopic" value of work-of-dissociation to exceed the "thermal" value by an amount equal to the energy of excitation of one of the resulting atoms, or perhaps to the sum of two values of energy of excitation, one corresponding to each atom and the other to the other. We need not therefore be at once alarmed if we find that the spectroscopic value is greater than the thermal value. The real question will be: whether the difference may reasonably be identified with an energy of excitation or a sum of energies of excitation of the forthcoming atoms?

The outstanding example of such an identification is probably that which was established by the several contributions of Franck, Dymond, Kuhn, and Turner, of whom the first three located the band-convergence-frequencies in the spectra of iodine, chlorine, and bromine, while the last made a spectroscopic determination of the energy-difference between the normal state and the adjacent metastable excited state of each of these types of atom. (The symbols for these states are 2P_2 and 2P_1 respectively.) The agreement seems good enough to justify the supposition that when any of these molecules is split apart by absorption of radiation, one of the forthcoming atoms is in its normal state, but the other is excited. Thus for iodine, the quantum which just dissociates the molecule has an energy of 2.47 equivalent volts; the *net* work of transformation of the molecule into a pair of normal atoms is evaluated from the thermal experiments as 1.50 equivalent volts; the excess of the former over the latter agrees very well with the value 0.94 which is the energy of the excited state of the iodine atom referred to the energy of the normal state as zero. Other cases of the sort are known; particularly interesting is that of sodium iodide, in which there is independent evidence that one at least of the atoms resulting from the dissociation is excited; for the vapor when irradiated by light of frequency

great enough to break the molecules apart, spontaneously emits the light (the *D*-lines) which is characteristic of sodium atoms passing from an excited to the normal state.

It is necessary now to take into account the fact that the stationary states of molecules must be classified (rotational fine structure being ignored, as heretofore) in a two-stage grouping: and to make this clear it is desirable to introduce the conventional molecule-model, even though in a few more years another model may supplant it.

We are, then, to visualize a diatomic molecule as a pair of massive particles—the nuclei of the two atoms—which float in a field of force so constituted that when their mutual distance d has a certain value, say d_0 , they are in equilibrium. When d is greater than d_0 , they are pulled together; when d is less than d_0 , they are pushed apart; when d is made slightly less or greater than d_0 , and the particles are thereupon left to themselves, they oscillate. The oscillations are not as a rule simple-harmonic, for the force acting on either nucleus is not as a rule strictly proportional to $(d - d_0)$. The various *vibrational states* correspond to various amplitudes of the vibration, which are determined by applying the quantum-condition when the field of force is known or postulated.

It is a sequence of such vibrational states, with vibrational quantum-numbers increasing by one unit from each to the next, which culminates in dissociation. There is however more than one such sequence. This we can foresee from the fact that the field of force is produced by the charges of the nuclei and those of the electrons jointly. Were the nuclei alone, the force would consist solely of the inverse-square repulsion which each exerts upon the other. The electrons contribute the additional force which at the distance d_0 just compensates this repulsion and at distances greater than d_0 overcomes it. Now in atoms, there are multitudes of different stationary states corresponding to different arrangements of the electrons. Molecules might be expected to resemble atoms in this respect; and as a matter of fact there are many kinds of molecules which are definitely known to exhibit several distinct arrangements of the electrons—several

distinct *electronic levels*, as I will call them in order to reserve the name "state" for the various states of vibration. For, as one easily sees, to each electronic level there belongs an entire sequence of "permitted" vibration-amplitudes; and the condition of the molecule is defined when and only when both the electronic level and the vibrational quantum-number are specified. In general, the permitted values of vibration-amplitude and of vibration-energy will not be altogether the same in the sequences associated with the different electronic levels, for they are influenced by the field of force, which changes when the arrangement of the electrons is changed.

Now let us return to the case of a converging sequence of absorption bands, such as occurs in the spectrum of iodine. Here the initial state common to all the transitions of which these lines are the tokens is the vibrational state of lowest quantum-number⁴ associated with the *normal* electronic level. The final states of these consecutive transitions are the vibrational states with consecutive quantum-numbers, associated with a particular *excited* electronic level. In effecting any of the intermediate steps which lead up to dissociation, and in effecting dissociation itself, the light not only increases the amplitude of the nuclear oscillations but also changes the arrangement of the circum-nuclear electrons.

How then are we to visualize the process of dissociation? According to a fertile idea first expressed by Franck, it is probably correct to regard the change in the arrangement of the electrons as the cause of the change in amplitude of oscillation, even when this is so great that the nuclei tear themselves apart altogether. When the electrons shift their places (or perhaps we should say, when the cloud of negative charge redistributes itself?) the nuclei suddenly find themselves moving with their former speeds in an entirely new field of force. Usually, though not by any means always, the equilibrium-distance d_0' in the new field is greater than the equilibrium-distance d_0 in the old. If then the nuclei are originally vibrating slightly or not at all in the

⁴ Sequences associated with vibrational states of higher quantum-numbers may also be observed.

vicinity of their former positions of equilibrium, they find themselves suddenly in a field which impels them outward; and perhaps before reaching their new positions of equilibrium they acquire so great a speed, as to dash right through these and completely out of the atom. This is dissociation. Or perhaps the speed which they acquire is not so great but that they are curbed and drawn back before they escape. In this case they continue to oscillate with an amplitude different from that which they possessed before; and this is excitation, disclosed by a line in the absorption-spectrum.

Why then, one may ask, does not always one of these things happen or always the other? Why does the rearrangement of the electrons sometimes result in dissociation and sometimes not? Probably the answer is, that even in the vibrational state of lowest quantum number, as well as the others, the nuclei are oscillating about their positions of equilibrium, so that the various molecules of a gas are at any given moment in all possible phases of vibration, and those in which the nuclei are closest together are dissociated because these particles are forced apart with great vehemence, while those in which the nuclei are farthest apart may hold together because for them the distance between the nuclei at the critical instant happens to be nearly the same as the new equilibrium-distance. It follows then that the fields of force before and after the rearrangement, and the amplitude of the vibrations before, jointly control the possibility or the relative probability of dissociation and mere change of amplitude of oscillation. Some spectra in fact reveal that both occur; in some, on the other hand, one sees only the continuous range of absorption but not a sequence of lines leading up to it; in yet others a sequence of lines fades out without reaching an obvious limit or the boundary of a zone of absorption. By means of the principle of correspondence, it is possible to draw inferences concerning the fields of force in the molecules which show spectra of any of these types.⁵

⁵ Such inferences are drawn likewise from observations on the relative intensities of the bands constituting a band-system in absorption or emission, under conditions where dissociation does not occur.

We have now considered cases in which dissociation results from the elevation, so to speak, of a molecule from its normal into an excited electronic level, this elevation being attended by so great an increase in vibrational energy of the nuclei that they tear themselves apart. There is yet another, and a very interesting, type of dissociation in which the atoms of a molecule tear themselves apart when the molecule spontaneously passes from an excited to the normal electronic level. This of course can occur only when the energy-difference between the excited and the normal level exceeds the work of dissociation—not an unusual condition among molecules, some of which can contain three or four times as much energy of excitation as would suffice, if properly administered, to disrupt them. An excellent example is furnished by the simplest of all molecules, that of hydrogen. Here we find a numerical agreement which goes far to confirm the general rightness of these interpretations of band-spectra.

In the absorption-spectrum of hydrogen there is a continuous region of strong absorption, the long-wave limit of which was located by Dieke and Hopfield at 850 \AA , corresponding to **14.53** equivalent volts. Beyond this limit lie sequences of bands—not one sequence, but two—corresponding to transitions from the normal arrangement of electrons into two distinct excited arrangements. Thus light of one frequency acting upon a normal molecule may set it into augmented vibration by changing the electron-system in one way, while light of another—perhaps even of quite the same—frequency may rearrange the electrons in quite another way and likewise alter the vibration of the nuclei. Either of these transformations of the electron system may be the prelude to dissociation, and dissociation when effected consumes in either case **14.53** equivalent volts of energy; this practically proves that the resulting fragments are in both cases the same.

Now this amount, **14.53**, is much in excess of the energy-difference between a pair of free normal atoms of hydrogen and a normal hydrogen molecule. This is the fact which was earlier interpreted as meaning that one at least of the atoms resulting from the dissociation is excited. At this point, however, the im-

portant feature is that the various excited states of the hydrogen molecule which are steps toward dissociation are themselves endowed with energy-values approaching 14.53, and much in excess of the energy sufficient to dissociate the molecule into two normal atoms. Does any intramolecular process ever occur whereby this amply sufficient amount of energy is actually applied to this purpose? If so, we may expect to find in the spectrum evidence of the following character:

Let us take as zero the energy of the hydrogen molecule in its normal electronic level and in its state of least possible vibration; and denote by B_n the energy of the molecule when in a certain excited electronic level (the one commonly designated as B) and in the state of n vibrational quanta. If then the molecule passes from the latter state to the former, the energy which it discharges is by definition equal to B_n . If instead it passes over not to the first but to the second vibrational state of the normal electronic level, the energy which it discharges is less than B_n ; if we denote by x_2 the energy of the second vibrational state of the normal level relatively to the first, then the energy in question is $(B_n - x_2)$. Similarly if we represent by $x_3, x_4, x_5 \dots$ the energy-values of the third, fourth, fifth \dots vibrational states of the normal level we shall have the sequence of energy-values $(B_n - x_3), (B_n - x_4), (B_n - x_5)$, and so on; and if finally this sequence of vibrational states converges at the energy-value x_∞ upon the state of dissociation, we shall have $(B_n - x_\infty)$ for the energy discharged by the excited molecule when it is dissociated upon attempting to return to the normal level. In the emission-spectrum, then, we shall find a sequence of emission-bands drawing closer together as the procession is followed from shorter wave-lengths to longer (the opposite sense from that prevailing in the sequences of which I have been writing); beyond these, a zone of continuous emission, the short-wave-limit of which, as the long-wave-limit of the sequence, yields the energy-value of the initial excited level *minus* the work of dissociation of the molecule. Now the energy-value of the initial level is itself given by the first band of the sequence. We have therefore only to subtract the limit-frequency from the frequency of the first

band of the sequence, in order to obtain the work of dissociation. Witmer made the requisite observations and computations for hydrogen, and obtained for the work of dissociation the value 4.34 equivalent volts.

Here we have an apparent disagreement between two values of work of dissociation obtained both by spectrum analysis; but the meaning is the same as that of the apparent discords between the spectroscopic and the thermal measurements on iodine and other vapors. *Dissociation of the hydrogen molecule leads in the two cases to two different conditions of the resulting atoms*—a statement which is established by the same sort of numerical agreement which I quoted in the case of iodine, only here a more accurate one, since both the values which are to be compared are derived from spectroscopic measurements. The question is, whether the difference between these two values can be identified with an energy of excitation of the hydrogen atom. Now the difference between the two values is $14.53 - 4.34 = 10.19$. The least excitation-energy of hydrogen is 10.15; and there is no other smaller than twelve volts, so that the coincidence can hardly be regarded as accidental. The almost inevitable inference is, that dissociation as the limit of the vibrational states of the normal level results in the formation of two normal neutral atoms, while dissociation as the limit of the vibrational states of the excited level *B* results in the formation of one normal and one excited atom. The “spectroscopic” value thus obtained for the heat of dissociation of the hydrogen molecule into normal hydrogen atoms far surpasses in accuracy that obtained by any other method.

The limitations of this article forbid more than a passing reference to the curious circumstance, which made the emission spectrum easy to interpret. In general, one would expect that hydrogen gas, when so violently stimulated as to emit light, would contain molecules in every vibrational state of the excited level,—a condition which would produce as many sequences of bands as there are of such states, and entail an extremely complicated spectrum. In this instance however the hydrogen was mixed with argon; and the excited atoms of argon possess just

the amount of energy which, when communicated in collisions of the second kind to normal hydrogen molecules, transfers them into one particular vibrational state of the excited level.

The most interesting question remaining is probably that of the laws, if any there be, which govern the nature of the particles into which a molecule is likely to fall apart. Could we, from the chemical or physical properties of a molecular gas, predict whether on being irradiated it would resolve itself into normal neutral atoms, into excited neutral atoms, or possibly even into ions?

The facts are rather surprising. There are certain molecules, of which the alkali halides and hydrogen chloride might be taken as examples, in which it seems likely on various grounds that an electron of one atom has been taken over into the system of the other. According to the well-known octet theory the sodium atom in the sodium chloride molecule, for instance, is practically ionized, having ceded its valence electron to the chlorine atom which thus becomes practically a negative ion. There are other molecules, however, for which there is no evident reason for expecting anything of the sort; as for instance those made of identical atoms, which might be expected to have perfect symmetry.

It might then be anticipated that molecules of the former type would split by preference into ions, or at any rate into highly excited atoms; while those of the latter sort would fall apart into pairs of normal neutral atoms. Well! the surprising thing is, that so far as any rule at all emerges from the data, it seems to be the contrary of this. Molecules composed of like atoms are dissociated by absorption of light (we are not speaking of what happens when an excited molecule starts back to its normal level) into one normal and one excited atom; so also are molecules such as carbon monoxide and nitric oxide. Like them we have such supposedly outstanding examples of the other type as sodium iodide and silver iodide. On the other hand, the alkali halides display continuous zones of absorption so located as to suggest that in addition to that commoner mode of dissociation,

they may separate themselves on absorbing light into normal neutral atoms. This apparent discrepancy between the facts and the expectations may however be removed by the remodeling of our expectations which the new forms of mechanics seem likely to bring about.

THE NEW QUANTUM MECHANICS

J. H. VAN VLECK

Department of Physics, University of Minnesota, Minneapolis¹

I. INTRODUCTION

The quantum theory has been revolutionized within the past three years by the development of a new quantum mechanics which is a far more comprehensive and satisfying theory than the original form developed by Bohr, Sommerfeld and others in 1913 and subsequent years. This remarkable new mechanics cannot be regarded as the product of any one man, but instead must be considered the result of the reaction of mind on mind among European talent in theoretical physics. This new quantum dynamics has, in fact, been developed in a great diversity of mathematical forms, which present a rather confusing array to the student beginning the subject. However, it must be emphasized that these various formulations, though different in mathematical structure, are in harmony with each other, and yield substantially equivalent results when applied to physical or chemical problems. The three main mathematical forms are the following: (1) the matrix theory of Born, Heisenberg, and Jordan, (2) Schroedinger's wave mechanics, (3) the so-called "transformation theory," based on kinematical indeterminism, developed by Dirac and Jordan, and interpreted by Heisenberg. Of these three formulations, the third is the most comprehensive, and includes the other two as special cases. In this paper I shall not endeavor to give the mathematical foundations of the quantum mechanics, as that would be too long a task. Instead I shall begin by explaining some of the philosophy and logic underlying the new theory and shall later survey some of its accomplishments. The philosophy may be summarized in the

¹ Present Address: Department of Physics, University of Wisconsin, Madison.

statement that at atomic distances our concepts of space and time must be revamped. In fact Heisenberg's epoch-making development of the matrix theory was spurred by Born's repeated emphasis to his colleagues at Göttingen that the reason the old quantum theory was then (1925) failing was that we were all too anxious to use the same concepts of space and time within the atom as in ordinary measurable large-scale events. Einstein and the relativists made us rescrutinize the space-time correlation at cosmic distances, and now the quantum theories bid us do this at the other extreme of size. Now, after all, the concepts of distance and time have a meaning only when we tell how they can be measured. This is very nicely emphasized in Bridgman's recent book, "The Logic of Modern Physics." At ordinary distances we determine lengths by means of measuring rods and time intervals by clocks. However, one cannot use a meter stick to measure the diameter of an atom, or an alarm clock to record when an electron is at the perihelion of its orbit. Consequently we must not be surprised if within the atom the correlation of space and time is something which cannot be visualized, and that models cannot be constructed with the same kind of mechanics as Henry Ford uses in designing an automobile. After all, within the atom there may be no geometry in the ordinary sense.

The goal of theoretical physics and chemistry must ever be to explain observable rather than unobservable phenomena. The presence in a theory of unobservable quantities or concepts, such as the "phase" or instantaneous position of the electron in its orbit in the old quantum theory, must be regarded as an element of weakness rather than strength in the theory. What the physicist observes about an atom is primarily its radiations, while the chemist measures its heats of combination, affinities, etc. We may say that we have a sound atomic theory when we have a set of a small number of mathematical postulates from which these observed things can be calculated correctly, even though it forces us to discard the usual space-time models. The new quantum mechanics is often characterized as a matrix theory. Why is this possible? Because what is measured about the

radiations of atoms may be tabulated in the form of a matrix. The term "matrix" may sound rather formidable to those not mathematically trained, but I think I can intimate some of the uses of matrices if I point out their rough similarity to something with which most readers are doubtless more familiar; namely, a baseball schedule. In a baseball schedule the entry in row 3 and column 2 gives the date when team 3 plays team 2. Now a typical quantum theory matrix is an expression of the form

$$\begin{vmatrix} A(11) & A(12)e^{2\pi i\nu(12)t} & A(13)e^{2\pi i\nu(13)t} & \dots\dots\dots \\ A(21)e^{2\pi i\nu(21)t} & A(22) & A(23)e^{2\pi i\nu(23)t} & \dots\dots\dots \\ A(31)e^{2\pi i\nu(31)t} & A(32)e^{2\pi i\nu(32)t} & A(33) & \dots\dots\dots \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \end{vmatrix}$$

Here the entry in row 3 and column 2, for instance, gives information about a transition between a 3 and 2 quantum state, just as the analogous baseball entry does about the meetings between teams 3 and 2. Each entry or "element" in the above matrix contains two factors; viz., an amplitude factor such as $A(32)$ and a frequency factor such as $e^{2\pi i\nu(32)t}$. The frequency factor determines the frequency $\nu(32)$ of the spectral line corresponding to the transition between the 3 and 2 quantum states. The amplitude factor determines whether such a line is intense, i.e., whether the transition probability is large or small. The intensity is, of course, proportional to the square of the amplitude; the exact formula for Einstein's transition probability coefficient $a_n \rightarrow m$ is $a_n \rightarrow m = 64\pi^4 e^2 \nu(nm)^3 |A(nm)|^2 / 3hc^3$. The diagonal elements, such as $A(11)$, $A(22)$, etc., give the properties which are associated with one stationary state rather than with the transitions between two states, and represent average values since the periodic exponential factors are absent. In the new quantum mechanics the Bohr frequency condition is still valid, so that we may still set $h\nu(nm) = W_n - W_m$. The conservation of energy means that the energy W is a "diagonal matrix" in which the periodic, or off-diagonal elements are all zero.

The above matrix scheme may be regarded as simply a method of recording laboratory observations regarding the intensity and frequency of spectral lines, for when we measure intensity we determine the amplitude A and when we measure the frequency

we determine ν . However, the crux of the theory is that the matrices can be calculated mathematically as well as measured experimentally, and wherever the difficulties of computation in the mathematics or of technique in the laboratory can be overcome there is excellent, in fact astounding, agreement between the theoretical and experimental matrices.

I shall not enter upon the matrix algebra and calculus with which it is possible to calculate the intensity-frequency matrices from the equations of motion, formulated of course in the matrix language. I shall simply mention that this matrix algebra is a non-commutative one, or in other words $px \neq xp$. Instead it is assumed that we have the quantum condition $px - xp = h/2\pi i$, where i denotes $\sqrt{-1}$, and h is Planck's constant, which thus enters as a measure of the "irrationality" of the theory, or peculiar non-commutativeness of the algebra. Here x is a coordinate of position, and p is its corresponding momentum. In the case of ordinary Cartesian coordinates, p is, of course, the product mv of mass and velocity. If we have such an extraordinary state of affairs as $xp = px$, then surely x and p cannot both be ordinary numerical quantities, for any ordinary numbers a and b of course possess the property that $ab = ba$. Instead p and x must be interpreted as matrices. From this Heisenberg and others conclude that *when dealing with amounts of action comparable with Planck's h , one cannot assign accurate numerical values simultaneously to a coordinate and its corresponding momentum. Thus within the atom it is meaningless to talk of a simultaneous position and velocity of the electron.* It is, of course, possible to attach a meaning to a simultaneous position and velocity of a large body, such as an automobile or even a dust particle, because in such we have an enormous number of quanta. Within the atom this is no longer the case. Experiments may be devised which will determine either x or p accurately, but not both x and p simultaneously. Instead if Δx be the error in specifying the position x , and Δp that in specifying the momentum p , then the product $\Delta x \Delta p$ of the two errors is always of the order of magnitude of Planck's constant, so that

$$\Delta x \Delta p \sim h \tag{1}$$

This is the so-called *Heisenberg indeterminism principle*, and, like the second law of thermodynamics, is very useful in predicting what experiments are possible, and what are inherently impossible. Thus high precision in position implies low precision in velocity, and vice versa, for the error Δv is of the order $h/m\Delta x$ and therefore increases as we decrease Δx . We may illustrate this in a crude arithmetical way by considering the motion of an electron. If we specify that this electron is exactly at a given position, say the origin, at a given instant $t = 0$, then the error Δx is zero, and consequently the error Δv is infinite; this means that the velocity is inherently undetermined and so can range anywhere from $-\infty$ to $+\infty$; so that at any subsequent instant of time, say one second, there is an overwhelming chance that the electron be an infinite distance away. Suppose, however, that instead of aiming to specify accurately the position of the electron at $t = 0$, we merely say that then it is somewhere between $x = -\frac{1}{2}$ and $x = +\frac{1}{2}$; then $\Delta x = 1$, and Δv is of the order h/m ; now $h = 6.55 \times 10^{-27}$, while for an electron $m = 0.9 \times 10^{-27}$ hence the error Δv in velocity is $6.55/0.9$, or about 7 centimeters a second. One could then not give accurately a numerical velocity, say 300 cm./sec., but one could say that the velocity lay somewhere between about $296\frac{1}{2}$ and $303\frac{1}{2}$ cm./sec., so that after 1 second the electron would be between about $x = 296$ and $x = 304$. (This is only a very crude calculation; a more accurate study leads to a Gaussian error curve.)

Another, and better illustration of the Heisenberg indeterminism principle is furnished by an attempt to measure simultaneously both the position and frequency of a packet or wave train of light energy. A prism will record a train of waves as being strictly monochromatic only if the train is infinitely long; the termination of the optical disturbance after a finite time interval will cause an interruption in periodicity and make the light in the prism appear diffuse. If the train contains n waves, it can be shown that the diffuseness in its spectrum will be of the order $1/n$, or in other words the fractional error $\Delta\nu/\nu$ in determining the frequency is comparable with $1/n$. If, following the light quant theory, we assign a packet or "quant" of light a momentum

$p = h\nu/c$, then Δp is of the order $h\nu/cn$. On the other hand the length of the wave train is $n\lambda$, where λ is the wave length, and consequently the position of the light is indefinite to an extent $\Delta x = n\lambda$. As $\nu = c/\lambda$, the product $\Delta x \Delta p$ of the two errors is of the order of Planck's h .

Because of the ambiguity resulting from the Heisenberg indeterminism principle, the future of a dynamical system can never be predicted with certainty. Instead only the *probability* that an electron be in a given configuration can be determined, and the future is only statistically determined. This is in sharp contrast to the "causality principle" of classical dynamics, whereby the subsequent history of a dynamical system is deter-

TABLE 1

CLASSICAL DETERMINISM	QUANTUM INDETERMINISM
Given initially x, y, z, v_x, v_y, v_z	Given v_x, v_y, v_z accurately; then x, y, z are undetermined (i.e., all values of x, y, z are equally probable), or vice versa x, y, z given and v_x, v_y, v_z undetermined
↓ (Arrow indicates calculation by classical dynamics)	↓ (Arrow indicates calculation by quantum dynamics)
Accurate determination of x, y, z, v_x, v_y, v_z at all subsequent times	Only statistical determination of x, y, z, v_x, v_y, v_z at subsequent times

mined if we know its initial coördinates and velocities. The uncertainty as to subsequent motion in quantum mechanics is however, perhaps not due so much to failure of the ordinary cause and effect relation as to the inevitably indeterminate character of the initial conditions, as by the Heisenberg indeterminism principle one cannot give both the initial position and velocity accurately. This contrast between classical determinism and quantum indeterminism should interest the philosopher as well as the chemist or physicist, and is indicated graphically in table 1. Thus quantum mechanics is essentially a means of calculating the probabilities of events. In a labora-

tory experiment one performs certain operations and observes certain consequences. The goal of a theory must ever be to explain "what goes out" in terms of "what is put in." At first thought it may appear as if the purely statistical correlation between cause and effect demanded by quantum mechanics is contradictory to the precision with which experiments can be performed. Experiments with large scale quantities, however, involve so many quanta that the Heisenberg indeterminism is

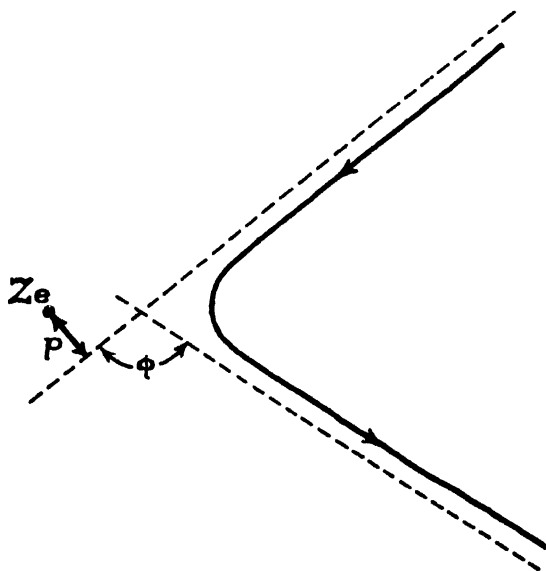


FIG. 1

obscured, and so there is only apparently a "sharp" correlation. On the other hand, our ordinary atomic or molecular experiments are in most cases fundamentally statistical in character, as what is measured is not particular values of the dynamical variables, but rather *average values* of certain functions of these variables, or else *distributions* telling how they are scattered over a wide range of values. Such statistical quantities have a meaning in and can be calculated with quantum mechanics, and so there is no contradiction of Heisenberg's principle. This idea is il-

illustrated in the next paragraph by a discussion of Rutherford's familiar experiment on the scattering of alpha particles by thin foils.

According to the classical theory of Rutherford's experiment, an alpha particle is deflected through a large angle in going through the gold foil if it happens to pass close to one of the gold atoms, and so is subject to large forces to bend it out of its course. On the basis of classical mechanics and the inverse square law it can be shown that the angle of deflection ϕ shown in figure 1 is connected with the closest distance of approach p which would result were the alpha particle to continue undisturbed in a straight line, by means of the relation

$$\cot^2 \frac{1}{2} \phi = \frac{p M v^2}{2 Z e^2} \quad (2)$$

Now assuming that all distances of approach are equally probable, or, more precisely, that the probability of the alpha particle approaching between distances p and dp is proportional to the area $2\pi p dp$ of an angular ring of inner and outer radii p and $p + dp$ respectively, it is often shown that the number dN of alpha particles scattered between angles ϕ and $\phi + d\phi$ is

$$dN = \frac{1}{4} \pi N n t \left(\frac{4 Z e^2}{M v^2} \right)^2 \operatorname{cosec}^2 \frac{1}{2} \phi \cot \frac{1}{2} \phi d\phi \quad (3)$$

where N is the number of alpha particles passing through the foil, t is the thickness of the latter, and Z is the atomic number of the atoms of the foil. n is the number of gold atoms per unit volume of foil, while M and v are the mass and velocity of an alpha particle. Formula (3) *which does not involve p* gives the so-called "distribution in angle" of the scattered alpha particles, and that is what is observed experimentally by counting scintillations on a fluorescent screen set at various angles of recoil.

Let us now turn to the quantum interpretation of the Rutherford scattering. Let us assume the initial beam of alpha particles is perpendicular to the foil, and that these particles have a common velocity. Then $v_x = v_y = 0$, while v_z equals v ,

supposing for simplicity that the face of the foil is the x - y plane. What does this mean?—that we are specifying all the velocity components of the alpha particles, and consequently by the Heisenberg indeterminism principle we cannot specify simultaneously the position coördinates x , y , z . Instead in a beam with definite velocity and direction the positions of the individual alpha particles within the beam must be regarded as entirely undetermined and it is hence meaningless to say whether an individual particle passes close to or far away from a scattering center (gold nucleus) of the foil. However, we have seen that the quantum mechanics determines the future motion of a particle in a statistical way. In the present problem this means that one can compute from quantum principles the probability that an alpha particle will recoil at a given angle. The “distribution in angle” of the scattered beam is thus determined, even though the concept of closest distance of approach p to the scattering center loses its meaning. This is in accord with experiment, for what is observed in the scintillation-counting is the relative abundance of different angles of recoil ϕ (cf. fig. 1), whereas there is no way of really measuring the quantity p .

The classical treatment of atomic problems introduces many concepts, such as the distance of approach p in the Rutherford experiment, which are not observed and hence not needed, whereas the quantum viewpoint usually divests the analysis of these superfluous quantities. This may be illustrated by the following rather foolish story. One time there was an Arabian chieftain who bequeathed to his oldest son one-half his camels, to his next oldest one-fourth of his camels, and to the third or youngest son one-sixth of them. There was great consternation in the tribe, because it was discovered at his death that he left only eleven camels, and fractional numbers of camels such as $11/2$, $11/4$ and $11/6$ were of no particular value. The matter was finally referred to an old sage of the tribe, who said, “I have only one camel, but am growing old, and have little more use for camels. I will therefore contribute my camel to the total if this will facilitate the apportionment.” This simplified things tremendously, for the fractions $12/2$, $12/4$, $12/6$ were all integers.

To the oldest son, for instance, the sage said, "Half of twelve camels is six. Take these six and be happy, for thou hast more than thy share." Similarly the two other sons were apportioned 3 and 2 camels, or more than their quota. But, behold, the beauty of the whole procedure is that $12/2 + 12/4 + 12/6 = 6 + 3 + 2 = 11$, so that the old sage received his camel back. Now the classical treatment of atomic problems revels in superfluous or non-observable variables which like the twelfth camel in this story are put into the analysis and then taken out again. The distance of approach, p , in the classical theory of the Rutherford scattering experiment is such an auxiliary camel, for although it is a variable put in the original equation (2), it is soon taken out of the analysis by assuming that all positions of approach are equally probable, and then integrating over all values of p from zero to infinity. This assumption is a statistical one which is not implied by the classical mechanics without supplementary hypotheses concerning probability, but is necessary to get the final distribution formula (3), which is what is actually verified in the laboratory. The quantum mechanics, on the other hand, has the advantage that it is fundamentally statistical in nature, and need not be supplemented by further assumptions concerning the probabilities of different values of quantities which cannot be observed. Instead the quantum treatment yields the final formula (3) without the intermediary of the microscopic, non-observable distance p .

Twentieth century developments, especially relativity and quantum mechanics, have forced the physicist to rescrutinize the real meaning of the variables entering in his equations and in so doing he has discovered that many customary concepts (e.g., absolute time, in cosmogony and instantaneous position of the electron in atomistics) which are as superfluous as the twelfth camel in the preceding story, and which have sprung up because one is too prone to visualize either cosmic or atomic space and time in terms of the same picture as for ordinary measurable distances. I wonder whether the chemist has not also suffered from the same kind of prejudices and likewise introduced concepts which prove to be hallucinations when put to close scrutiny.

The chemist has often thought that he has found strong evidence on the instantaneous positions of electrons, and first and last there has been considerable controversy as to whether the facts of organic chemistry can be explained as well by assuming that the electrons are moving, as the physicist would like, as by supposing that they are standing still. However, is it really necessary to suppose that the structural bonds of the organic chemists represent instantaneous positions of the electrons, or would it not do just as well to suppose that they represent average positions, for the mathematics indicates that it is the average rather than instantaneous positions of the electrons which determine whether the nuclei are in equilibrium? Or going a step further from a picture, would it not perhaps do to suppose that these structural diagrams are simply a way of indicating diagrammatically some of the symmetry properties of the solutions of Schroedinger's wave equation, to be discussed later? Recent work of London and Heitler seems to indicate that the systematization of chemical compounds is closely related to the group theory of mathematicians. However, one must not necessarily infer from this that the group characteristics are the geometrical characteristics. There is, of course, much evidence for tetrahedral models of the carbon atom, but does this necessarily mean that the instantaneous positions of the electrons project out in four directions; could not this evidence mean that this symmetry is only true of average positions of the electrons, or even that there is some dynamical function, important for the mechanics of chemical combination, but without any immediate geometrical significance, which is symmetrical mathematically in the variables representing the coördinates of the four electrons?

According to the theoretical physicist, one has no right to speak of the *instantaneous* position of the electron in its path within the atom, for if one knew where the electron were located each successive instant of time, one would know both the position and velocity of the electron, which we have seen is contrary to the Heisenberg indeterminism principle. It can, however, be shown that it is legitimate to introduce the concept of the *average* position of the electron; such averages are, in fact, given

by the diagonal elements of properly chosen Heisenberg matrices. It is these average positions, and not the instantaneous ones, which are revealed by experiments on the scattering of x-rays of the type which Professor Jauncey has so interestingly discussed in another paper of this symposium. Clearly the time of scattering of an x-ray quantum cannot be measured with any accuracy compared to the average time required for the electron to move from one side of the atom to the other. Chemists have often asked me this question, "Where are the electrons located in, say, the bond between two hydrogen atoms to form a hydrogen molecule. Does each electron remain with its own nucleus, or

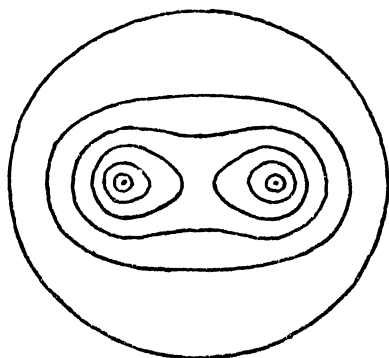


FIG 2 ATTRACTION

are the two electrons shared 50-50 by both nuclei." According to the quantum mechanics, the latter alternative comes the closer to the true state of affairs, for in the hydrogen molecule the average electronic charge distribution is symmetrical with respect to the two nuclei, and the two electrons are continually exchanging places, so that it is impossible to say which electron belongs with which nucleus. When the hydrogen atoms are too far apart to form a molecule, say at a distance 3×10^{-7} cm. comparable with the distances of approach between molecules in the kinetic theory of gases, the exchange of places is very infrequent, only about once in 10^{30} years on the average, whereas when the atoms are so closely knit together as to form a molecule,

the exchange transpires on the average about 10^{10} times per second! The statistical or average charge distribution in the hydrogen molecule can be calculated directly from the quantum postulates, and is shown in figure 2, taken from a paper by London. The contour lines in these figures represent the statistical charge density. Thus this density is large where the lines are close together, and small where they are far apart. The densest region is, of course, in general near the two nuclei. The mathematical analysis shows that there are two solutions of the Schroedinger wave equation corresponding to the interaction between two hydrogen atoms. In one of them, shown in figure 3, the forces exerted on the nuclei are entirely of a repulsive nature. The meaning of this is that two hydrogen atoms may be brought into contact without necessarily forming a molecule.

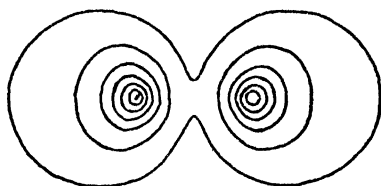


FIG. 3 REPULSION

In the other solution, figure 2, there is attraction as well as repulsion, and this is what makes possible the formation of a stable molecule. It is seen from figure 2 that in this other solution the negative electronic charge tends to pile up between the two nuclei, and it is the attraction of this negative charge which tends to bring the nuclei together and offsets the purely repulsive forces which exist between two positive charges by themselves. In the solution shown in figure 2, the nuclei are in positions of equilibrium, whereas in figure 3 they will fly apart.

The concept of a statistical or average charge density, such as is shown in figures 2 and 3, can perhaps be rendered more graphic by the following comparison. Supposing one were to photograph a swarm of fireflies at night by means of an exceedingly sensitive camera. If one were to take an instantaneous photo-

graph or "snap-shot," each firefly would appear as a bright point on the photographic plate. If we took snap-shots at frequent intervals we could then trace the motion of each of the flies. Suppose, however, that instead we were to take a time exposure extending over a long period. Then the luminous points would be blurred out into a cloud of light upon the plate. The cloud would be brightest where the fireflies are most apt to congregate. Now the information which it is possible to obtain about the motion of the electron within the atom in quantum mechanics is analogous to that given by the time exposure in the firefly illustration. The statistical charge density corresponds to the density of the luminous cloud. Efforts have sometimes been made to interpret the electron in quantum mechanics as itself a sort of nebulous body spread over the entire atom, but that is no more correct than to say that a firefly is a bright cloud extending over a large distance. In either case it is only the time average of position that has the cloud or fluid-like appearance. The essence of the Heisenberg indeterminism principle is that a "snap-shot" of the electronic motion is inherently impossible, for position cannot be instantaneously specified if the energy has a definite value corresponding to a stationary state of the atom. It appears almost paradoxical that although one can never say *when* an electron is in a given position within the atom, one can nevertheless talk of the fraction of the time that it is in this position, for clearly the places of large average charge density are those where the electron spends a large fraction of its time.

II. THE WAVE VERSION

The new quantum mechanics is often spoken of as a "wave mechanics." Now the concept of waves has certainly not appeared in any of the material which I have presented so far, and so one may be wondering where this quantum theory acquires an undulatory nature. In order to introduce the wave idea it is necessary to probe a little further into the mathematical formulation of the theory. In this there is involved a function $\psi(x, y, z)$ which has the property that $|\psi(x, y, z)|^2 dx dy dz$ measures the probability that the electron be in a small element of volume

$dx dy dz$. Thus the square of the magnitude of this function ψ may be regarded as proportional to the statistical charge density which we have already discussed at length. The places where the electron is likely to be located are those where ψ is large, while those where ψ is numerically small are seldom visited. The function ψ is determined by solving the differential equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{8\pi^2 m}{h^2} (W - V) \psi = 0 \quad (4)$$

Here M is the mass of the particle, $V(x, y, z)$ is its potential energy in a field of force, and W is its total energy, a constant quantity. This is the Schroedinger wave equation, which although only two years old, is one of the most celebrated equations of mathematical physics. Its solution $\psi(x, y, z)$ which determines the statistical charge density, is called the "Schroedinger wave function." The term "wave" is used, because in many problems it is found that ψ exhibits wave-like recrudescences of maxima and minima. I shall not attempt to give the mathematics which Schroedinger used in getting this equation. Following a cue given by suggestive work of de Broglie, he found that this equation could be obtained by assuming that the modifications which it is necessary to introduce into the equations of classical mechanics to get those of quantum mechanics are very similar to those which it is necessary to introduce in the theory of geometrical optics to get that of physical optics. Now geometrical optics does not reveal the undulatory nature of light, but can usefully be employed as long as the optical objects are large compared to the wave-length, and we do not probe into the interference and diffraction phenomena in which the wave effects appear. Thus x-rays have such short wave-lengths that most experiments on their propagation can be described by geometrical optics, as it is a sensitive and difficult experiment to show the diffraction of x-rays. Now Schroedinger held that ultimately mechanics, like optics, has a wave-like structure, and that the reason ordinary large-scale experiments can be described by classical mechanics is because all the objects are large compared to the wave-length. On the other hand, in atomic

phenomena the diameter of atoms is comparable with the quantum wave-length, and so it is necessary to use a new quantum mechanics which is the analog of physical rather than geometrical optics. This optical analog is exceedingly alluring, and it is dramatic to say that "Schroedinger has done for mechanics what Fresnel has for optics." However, it must be mentioned that it is possible to derive the equation (4) from the matrix theory without any appeal to the optical analog or the wave idea. The correlation with the matrix theory is, in fact, perhaps the most rigorous way of deriving (4). At first it seems like a far cry from the wave viewpoint to the Heisenberg matrices which we introduced near the beginning of our discussion. However, it can be shown that all the matrix elements can be determined by evaluating certain integrals involving the wave function ψ . In particular, if $x(nm)$ be the matrix elements of a coordinate of position x , it can be shown that

$$x(nm) = \iiint \psi_n^* x \psi_m dx dy dz, \quad (5)$$

where the integration extends throughout all space, while ψ_n and ψ_m are respectively the wave functions belonging to the stationary states n and m . ψ^* means the complex conjugate of ψ . One way of interpreting the wave function ψ is that it is simply a mathematical tool for calculating the matrix elements by means of equation (5). This formal mathematical interpretation of ψ , however, does not give ψ the physical significance which it deserves. At the same time it must be mentioned that while Schroedinger has at various times expressed the hope that his wave mechanics would remove the "irrational" or discrete characteristics of our atomic dynamics and make it possible to construct an essentially classical model of the atom without the rather forced introduction of sudden quantum jumps, the consensus of opinion of theoretical physicists today seems to be that we cannot keep our classical pictures in dealing with small scale phenomena, and that instead the language of probabilities must be used with all the indeterminism involved in Heisenberg's principle.

Now phenomena described by means of differential equations

in general exhibit a continuous rather than discrete character, for all of nineteenth century mathematical physics, with its continuous media, etc., was constructed primarily on differential equations. Consequently one may at first wonder how quantum discreteness can come out of formulation of atomic dynamics by means of differential equations. The answer is that we must admit only solutions of (4) which are of what I may dub of a "civilized character," and such solutions in many instances form a discrete succession even though in general a differential equation admits a continuous infinitude of solutions. By a "civilized solution" I mean one in which the wave function ψ is continuous, single-valued, and finite throughout all space. Most solutions do not possess this property. Very often, for instance, ψ increases without limit as we go out to infinity, and such a solution is worthless, as it would mean that the statistical charge density is greatest at infinity, and that the electron is infinitely far away most of the time. Instead it can be shown that there will be a civilized solution only if certain particular values be assigned to the constant W appearing in equation (4). Such values are called "Eigenwerte" or "characteristic values" by mathematicians, and the corresponding solutions ψ are the "Eigenfunktionen" or "characteristic functions." Schroedinger's papers bear the title "Quantisierung als Eigenwertproblem." These "Eigenwerte" are the values of the W 's to be used in the Bohr frequency condition $h\nu = W_1 - W_2$, and thus *the quantized values of the energy are those which make Schroedinger's wave equation have a civilized solution.* Sometimes these allowed values of W form a continuous rather than discrete succession, and the quantum mechanics is thus able to explain very elegantly, directly from the fundamental postulates, the existence of both continuous and line spectra. When, for instance, the electron is completely liberated from a hydrogen atom, it can be shown that the "Eigenwerte" are continuous, and this means that when a free electron is captured by a hydrogen nucleus a continuous spectrum is emitted, whereas a line spectrum is emitted when there is a transition of a bound electron between two stationary states. This is in accord with the fact that beyond

the "convergence frequency" of spectral series, corresponding to ionization of an atom, a continuous spectrum is observed experimentally.

How the requirement of a civilized solution ψ quantizes the energy may be illustrated by considering a particularly simplified dynamical system; viz., a dipole of moment of inertia I constrained to rotate in a plane. If φ denote the angle through which the dipole rotates, it can be shown that here the wave equation reduces to

$$\frac{\partial^2 \psi}{\partial \varphi^2} + \frac{8\pi^2 I}{h^2} W \psi = 0$$

This is a differential equation of the same type form as that of simple harmonic motion, and has the solution

$$\psi = A \cos [(8\pi^2 IW/h^2)^{\frac{1}{2}} \varphi - \epsilon]$$

where A and ϵ are constants of integration. Now to be "civilized," the function ψ must be single-valued, and this means that ψ must revert to its original value when we go once around a circle and return to the starting point. Hence we must have $\psi(\varphi + 2\pi) = \psi(\varphi)$. Now a function of the form $f(\varphi) = A \cos(k\varphi - \epsilon)$ possesses the property that $f(\varphi + 2\pi) = f(\varphi)$ only if k is an integer n , as the cosine reverts to its original value when its argument increases by an integral multiple of 2π . Hence, as in our problem $k = (8\pi^2 IW/h^2)^{\frac{1}{2}}$, we must have $W = n^2 h^2 / 8\pi^2 I$, where n is an integer. This is nothing other than the familiar so-called Deslandres formula for the energy of a rotating dipole in the quantum theory,—a formula which is constantly being used for describing rotational energy levels in molecular spectra. It should be mentioned that as we have derived the formula, the quantum number n is a whole integer, whereas to secure accord with experiment it is necessary for n to be a "half quantum number." However, when the rotating dipole is treated as a three- rather than two-dimensional problem, as of course it must be, the formula with half quantum numbers comes correctly out of the new mechanics, whereas it did not in the old quantum theory.

The mathematics which is used in solving Schroedinger's equation shows a rather striking similarity to that used in solving many problems in what I may term "nineteenth century mathematical physics"—problems in elasticity, vibrations of ropes and membranes, and the like. In fact many of the equations have already been solved before advent of the quantum theory by workers in elasticity, etc., and a very nice resumé of such classical "Eigenwertprobleme" is given in Courant and Hilbert's "Methoden der Mathematischen Physik," a book which is constantly being quoted in papers on the new quantum theory. For this reason, the wave formulation of the new mechanics is apt to appeal most strongly to mathematical physicists who have been trained primarily along the line of classical nineteenth century mathematical physics, whereas the matrix formulation appeals more strongly to those who have been trained in the mathematics of the old quantum theory, especially the correspondence principle. In classical vibration problems the solutions of the differential equations must fulfill certain so-called "boundary conditions"—the displacement, for instance, must vanish at the ends of a rope if it is fastened at either end. The differential equation for a uniform rope is

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2}$$

where v is a constant. If now we assume a periodic solution $y(x, t) = u(x) \cos(2\pi\nu t - \epsilon)$, then

$$d^2u/dx^2 = -4\pi^2\nu^2v^{-2}u$$

The solution of this equation is

$$u = A \cos(2\pi\nu x/v) + B \sin(2\pi\nu x/v)$$

Let l be the length of the rope. Since the rope is fastened at either end, the displacement y must vanish at $x = 0$ and also at $x = l$. The first of these two conditions shows immediately that A equals zero, as the cosine does not vanish at $x = 0$. The vanishing of y at $x = l$ shows that $\sin(2\pi\nu l/v) = 0$ and hence $2\pi\nu l/v = n\pi$, where n is an integer. Thus the possible

vibration frequencies of the rope are $\nu = nv/2l$ ($n = 1, 2, \dots$), and form a definite discrete succession (fundamental, first harmonic, second harmonic, etc.). These are the "normal modes of vibration," and the general motion is the superposition of all of them. The important thing to be noted in this classical rope problem is that the boundary condition that the displacement vanish at the two ends restricts the vibration frequencies to certain particular values. Now in much the same way in quantum mechanics the energy and hence the spectroscopic frequencies given by the Bohr frequency condition are restricted to certain particular values by the requirement that ψ be a civilized function throughout all space. The latter requirement can be regarded as a boundary condition, for mathematicians show that at singular points of a differential equation, where the general solution ceases to be analytic, the requirement that ψ be merely finite, without specifying its numerical value, is as much of a restriction as can be imposed at these points, and so can be regarded as a boundary condition at such points. Thus it is in a sense the boundary conditions which limit the number of appropriate solutions of Schroedinger's equation and restrict the energy to certain particular quantized values. Hence quantization is formally similar to a classical boundary value problem. It is tempting to say that the characteristic or Schroedinger wave functions represent the normal modes of vibration of the atom, just as do the various harmonics in the rope problem, but this is a bit misleading, as the spectroscopic frequencies are proportional to the difference of two energy values, and so always involve a pair rather than just one stationary state or wave function.

A particularly simple case of the Schroedinger wave equation is that which corresponds to an electron moving in a straight line, say the x -axis, in the absence of any external field of force. If we set the potential energy V equal to zero, and assume that ψ is a function only of x , the wave equation (4) reduces to

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2mW}{h^2} \psi = 0$$

The solution of this is

$$\psi = A \cos \left[(8\pi^2 m W / h^2)^{\frac{1}{2}} x - \epsilon \right]$$

or if we set $W = \frac{1}{2}mv^2$, we can write this as $\psi = A \cos [2\pi m v x / h - \epsilon]$. Now the expression for a standing² wave of wave-length λ is $A \cos [2\pi x / \lambda - \epsilon]$. Hence the solution of the wave equation can be regarded as a plane wave of wave-length

$$\lambda = h / m v \quad (6)$$

That an electron beam has a wave-length of the value (6) is shown very nicely in the already celebrated recent experiments of Davisson and Germer. They show that when an electron beam is scattered by reflection from a single crystal of nickel, the spatial distribution of the scattered electrons is the same as that of the scattered x-rays in Laue's or Bragg's experiments on the scattering or reflection of x-rays from a crystal, provided the wave-length of the x-rays is assumed to have the value given by (6). Thus the electron scattering resembles optical scattering for a wave-length $\lambda = h / m v$. In general the quantum mathematics of scattering of an electron beam by impact with atoms is very similar to a classical diffraction problem in optics. Hence there is a good deal of talk at present about "waves of matter," and it is even said that the electron is a wave. This, however, is not quite right. It is much better to say that instead we have waves of probability which govern the motion of the electron, for we have seen that the Schroedinger wave function determines a statistical charge density, and the directions in which the scattered or diffracted wave has a large amplitude are those in which the electron is likely to be scattered. The diffracted wave, like a ripple, extends out in all directions, whereas an

² The Schroedinger waves appear here to be of the standing rather than moving type, because we have not used a rather more general form of his equation which allows for explicit appearance of the time, and in which ψ is a function of x, y, z, t , rather than just x, y, z . The dependence on the time in a conservative system proves to be of through a purely exponential "time factor," $e^{2\pi i W t / h}$ and makes the waves moving rather than stationary. The absolute value of the time factor is unity, and so does not affect the value of the statistical charge density.

individual electron can, of course, recoil in only one direction. The wave amplitudes simply determine the angles in which it is likely to go.

The Davisson-Germer experiments and the theory of quantum mechanics show that the motion of matter has undulatory as well as corpuscular aspects. Conversely the Compton effect shows that in some respects radiation exhibits a corpuscular nature rather than the exclusively undulatory nature presupposed in the wave theory of light. Thus matter and radiation both exhibit at the same time the traits of corpuscles and of waves. It has often appeared as if there were a flat contradiction between the corpuscular and wave viewpoints, especially in the case of radiation. However, the statistical interpretation of motions demanded by quantum mechanics at last removes this dilemma, and to a large extent the motions of particles, either of matter or of radiation, can be regarded as guided by waves of probability. This situation is reminiscent of the old theory of Bohr, Kramers, and Slater, but with the difference that the laws of conservation of energy and momentum now hold accurately, a possibility later suggested by Slater himself. The work of Dirac, in particular, is very important in showing how the properties of the electromagnetic field can be derived from quantum mechanics and statistics.

If the motion of *matter* is attended by wave-like maxima and minima, it seems at first a little surprising that this fact was not unearthed experimentally before the recent experiments of Davisson and Germer, whereas the wave nature of *light* has been known for over a century. Why should not, for instance, the quantum wave length h/mv appear in the classical experiments of Rutherford on the scattering of alpha particles? The explanation is that since the wave-length is inversely proportional to the mass, it is about 8000 times smaller in the scattering of alpha particles than in that of electrons. Hence in Rutherford's experiment the wave maxima and minima lie too close together to have been detected, and so it is possible to describe the experiment by a mechanics analogous to geometrical optics, in other words classical mechanics. Because of the short wave length the

classical result has had the good fortune to agree with the quantum one, and so could explain Rutherford's experimental results, whereas in the Davisson-Germer experiment this agreement disappears and the classical theory breaks down.

III. WHAT THE QUANTUM MECHANICS HAS ACCOMPLISHED FOR THE PHYSICIST

I shall now turn abruptly from this rather sketchy discussion of some of the underlying ideas of the new quantum theory to a recital of some of the things which it has accomplished in the way of applications. As this new dynamics has been developed primarily by physicists rather than chemists, most of the applications are naturally to physical rather than chemical problems. Considering that the new quantum mechanics is only celebrating its third birthday, the success which it has had in "clearing up" problems that baffled the old quantum theory is little short of miraculous. One reason for this triumph is, of course, the large number of physicists that have been engaged in working over well-known material with the aid of the new theory. Ten of the items which the new quantum mechanics has to its credit are the following:

1. *Everything that the old quantum theory did correctly.* Although the old quantum theory had many inadequacies, it did nevertheless furnish a clue to the interpretation of an enormous mass of material, especially spectroscopic. The old theory was noted, perhaps more than for any one thing, for quantitative explanation of the spectra of hydrogen and ionized helium. Exactly the same numerical formulas, notably the Balmer formula for spectral series, Sommerfeld's fine-structure formula and the Epstein-Schwarzschild expression for the Stark effect, are obtained with the new theory as with the old. In non-hydrogenic atoms the presence of a large number of electrons usually makes it impossible to calculate quantitatively the position of spectral lines. But we all know that in the old quantum theory the concepts of stationary states, and of the Bohr frequency condition with its attendant Ritz combination principle, were exceedingly fruitful in explaining qualitatively the facts of non-

hydrogenic spectra, and in unravelling a maze of spectroscopic data which it would have been hopeless to correlate otherwise. All this can be done equally well with the new mechanics. Although we cannot attach as much reality and vividness to electron orbits in stationary states as previously, nevertheless the new theory has the vital feature of a discrete succession of energy levels, with spectroscopic frequencies proportional to their differences. Thus we can still keep our beloved energy level diagrams, selection principles, and the like. In particular, the fundamental relation $Ve = h\nu$ correlating critical potentials V with spectroscopic frequencies ν still remains valid.

2. *Fine details of the hydrogen and ionized helium spectrum not explained previously.* Although the old quantum theory gave in the main a very successful explanation of the hydrogen spectrum, it became increasingly obvious that in certain minor respects the old Bohr theory was in pretty flat contradiction with experiment. Among these may be mentioned the observation of a so-called "Paschen-Back" effect, or merging of the fine-structure in an extremely intense magnetic field, and the detection of certain fine-structure components which violated the selection principle. Also in some cases, especially in the Stark effect, the old theory gave more rather than fewer lines than were necessary to explain the experimental results, and to get rid of superfluous components it was necessary to rule out rather arbitrarily certain states on the ground that in them the electron would collide with the nucleus. All these difficulties disappear with the advent of the new quantum mechanics.

3. *The spectrum of neutral helium.* Although the old quantum theory dealt with ionized helium pretty successfully except for the items (2), the neutral helium atom, on the other hand, was a hopeless stumbling block to it. Dozens of models were proposed for the arrangement of the electron orbits in neutral helium, but none agreed with the observed critical potentials and optical frequencies. All this is now changed, for Heisenberg has calculated the term values of the excited states of helium with the new mechanics, while Slater, Kellner, and Sugiura have all independently computed the energy of the normal state. The agree-

ment which they find with experiment is gratifying when it is considered that the calculations can only be made approximately because in neutral helium we have all the mathematical difficulties of the celebrated three-body problem.

4. *A theory of intensities as well as frequencies.* The old quantum theory was one-sided. It gave a perfectly definite scheme for calculating the frequencies of spectral lines, but it furnished no means of determining quantitatively their intensities, although actually the two things must be connected. In the old theory we could gauge qualitatively the intensity of spectral lines by appealing to the correspondence principle, but this borrowing of the results of classical theory lacked adequate logic, and never gave unambiguous numerical results except in the particular case that it showed certain lines should be of zero intensity, or in other words that certain transitions should not occur. These exclusions of certain transitions gave selection principles of great usefulness, and the crude theory was adequate to predict zero intensities even though in general it might be in error by several hundred per cent, because fortunately any percentage of zero is still zero. The new mechanics not only gives the selection principles without artificial appeal to the correspondence principle, but also permits a "refined" or accurate calculation of the brightness of lines whose intensities are different from zero. In the mathematics of the matrix theory the intensities and frequencies are beautifully woven together. We have, in fact, mentioned near the beginning of the paper how the Heisenberg matrices contain both amplitude and frequency factors.

5. *The anomalous Zeeman effect.* It is found experimentally that in a magnetic field H the "Zeeman" displacements of spectral lines are usually equal to fractional multiples of the normal Lorentz unit $He/4\pi mc$, rather than to this unit itself. This is the phenomenon of the "anomalous Zeeman effect." Any ordinary theory would give only the normal effect, and the old quantum theory could explain the anomalous effect only in semiempirical fashion and with the aid of many contortions. In the new theory, of course, the anomalous effect comes out

naturally. If I had been giving this paper just two or three months ago I would have said that the new quantum mechanics could explain the anomalous Zeeman effect only with the aid of the *spin electron*. By the term spin electron is meant the idea that the electron behaves like a spinning top, and so has an internal or rotational degree of freedom besides its three translational degrees of freedom. Such a fourth degree of freedom and hence fourth quantum number appears definitely to be demanded by the Smith-Stoner-Pauli classification of the periodic table. The suggestion that the electron have a magnetic moment due to internal motion is by no means a new one, and has been made by Parson, Compton, Kennard, and others but its incorporation in the spectroscopic work of quantum theory is due primarily to Uhlenbeck and Goudsmit. For a long time it appeared as though the internal spin was a hypothesis entirely supplementary to the quantum mechanics. However, within the past month or two, brilliant work by Dirac has shown that the requirement that the Schroedinger wave equation be orthogonal and compatible with relativity (which apparently demands several simultaneous wave equations) is alone sufficient to yield the dynamical terms ordinarily attributed to electronic spins. Thus in a certain sense the spin electron is the child of the new quantum mechanics rather than a supplementary assumption, but it is a bit doubtful if it would ever have been discovered if approached originally from the abstract mathematical rather than empirical spectroscopic viewpoint. The inclusion of the spinning motion of the electron is very vital in the theory of any magnetic effect, such as the Zeeman effect, or magnetic susceptibilities (item 7 below), as the internal motion gives the electron a magnetic moment even when its center of gravity is at rest. The spin is even necessary in interpreting the hydrogen fine-structure with the new mechanics. It also underlies the so-called "gyromagnetic anomaly;" viz., that the experiments of Barnett, de Haas and others on rotation by magnetization, or the converse, reveal a ratio of magnetic moment to angular momentum of twice the classical value. This is intimately connected with the fact that the ratio of the internal magnetic

moment to internal angular momentum of the electron is $-e/mc$, or twice the value for ordinary orbital motions. This factor two is what makes all the anomalies in the Zeeman effect, gyro-magnetic effect, etc.

Both in x-rays and in visible spectral regions there are certain doublet lines whose origin was previously shrouded in mystery. The doublet intervals obeyed numerically Sommerfeld's relativity formula, but nevertheless the doubling could not be due to relativity corrections. They are now interpreted as "spin doublets" due to two different orientations of the spin axis relative to the rest of the atom.

6. *Half quantum numbers no longer a bugaboo.* To secure agreement with experiment in the old quantum theory it was necessary to give certain quantum numbers half integral values $1/2, 3/2, \dots$ rather than the integral values 0, 1, 2. Such a procedure was most disconcerting, as the rigorous old theory demanded only whole integers. The new mechanics, however, very obligingly gives unambiguously half quantum numbers where they are needed and whole quanta where the latter are required.

7. *Improvement in the theory of dielectric constants and magnetic susceptibilities.* The preceding items have been primarily spectroscopic in nature, and often deal with excited rather than normal states. The chemist is apt to conceive of the physicist as some one who is so entranced in spectral lines that he closes his eyes to other phenomena. However, it must be mentioned that a theory of spectral lines is also one of critical potentials. Also in the theory of dielectric constants and magnetic susceptibilities we have another quite different field in which we can test the quantum mechanics, and which involves only the normal, in distinction from excited, states of molecules. In this field the new quantum theory is far more successful than the old one. One reason for this is that the new mechanics gives results in much closer accord with the classical theory than did the old quantum theory. In the latter there was a conflict with classical theory even at high temperatures, and Pauling even showed that an innocent little magnetic field of 100 Gauss or so would have a prodigious effect on the dielectric constant, quite contrary

to experiment. In the new mechanics or in the pure classical theory this prodigious effect disappears. The new theory restores the factor $1/3$ very generally in the Langevin-Debye formula. Because of the changes in the constant in this formula with the gradual development of quantum theory, there have been corresponding alterations in the electrical moments deduced from experimental data on dielectric constants. The electrical moment of the HCl molecule, for instance, has had quite a history, as indicated in table 2. The first column gives the value of the constant C in the Langevin-Debye formula:

$$(\epsilon - 1)/4\pi = N\alpha + (NC\mu^2/3kT)$$

Thus fortunately the electrical moment of the HCl molecule reverts to its classical 1912 value. The quantum theory of

TABLE 2

VALUE OF CONSTANT C	FORM AND YEAR OF THEORY	CORRESPONDING VALUE OF ELECTRICAL MOMENT μ OF HCl MOLECULE
$\frac{1}{3}$	Classical, 1912	1.034×10^{-18} e s u.
1.54	Whole quanta, 1921	0.492×10^{-18}
4.57	Half quanta, 1925	0.322×10^{-18}
$\frac{1}{3}$	New mechanics, 1926	1.034×10^{-18}

susceptibilities, however, does more than merely agree with classical theory regarding the factor $1/3$. Since magnetic moments are quantized, it permits the calculation of the absolute value of the magnetic susceptibilities whenever the spectroscopic term values are known. The only common paramagnetic gases are NO and O₂, and the numerical magnitudes of their calculated susceptibilities agree with experiment to within 1.5 per cent.

All the preceding results can usually be obtained with the ordinary matrix or wave formulations of the quantum mechanics. The remaining three items, however, mostly require the use of the more general "transformation theory" of Dirac and Jordan, or its equivalent.

8. *Dispersion and the emission and absorption of radiation.* Dispersion was particularly bothersome in the old quantum

theory, which could never explain why the resonance frequencies in dispersion were experimentally the spectroscopic frequencies given by the Bohr frequency condition rather than the altogether different frequencies of motion in the orbits constituting the stationary states. The new mechanics, however, yields the Kramers dispersion formula, previously derived semi-empirically from the correspondence principle. In the new theory the dielectric constant equals the extrapolation of the square of the index of refraction to infinite wave lengths, as it should, whereas the validity of this relation was doubtful if the Kramers formula was forced into the old theory. As the result of masterful treatment by Dirac, a mechanism has at last to a certain extent been found for the previously so mysterious quantum jumps between stationary states. Dirac finds it is possible to treat radiation as a dynamical system obeying the Einstein-Bose statistics, and starting from such a postulate he is able to derive Einstein's values for the "A" and "B" transition probability coefficients. Dirac's work brings out nicely the parallelism between matter and radiation, and their corpuscular and wave aspects, which are complementary rather than contradictory.

9. Collisions of electrons with atoms. The old quantum theory furnished no real mechanics governing electron impact phenomena although a rather blind conservation of energy argument furnished the fundamental relation $Ve = h\nu$ between critical potentials and frequencies. The theory of collision phenomena has been developed with the new mechanics by Born and others. As already mentioned, there is a good deal of similarity to a diffraction problem in optics, and the square of the amplitude of the diffracted wave measures the probability that an electron will be scattered at a given angle. The requirement that the solutions be "civilized" yields nicely the proper discrete values for the critical potentials. In the case of collisions with hydrogen atoms, Born is able to compute numerically the excitation and ionization efficiencies, the angular distribution of scattered electrons, etc. The wave-like maxima and minima in the latter are confirmed at least qualitatively in the experiments of Dymond on the scattering of electrons in helium. A closely allied question

is the polarization of radiation excited by electron impact, and the theory of this has been developed by Oppenheimer. A most peculiar phenomenon is the so-called "Ramsauer effect," whereby slow-speed electrons are able to go right through inert gas atoms with very little scattering. The theory of this has at last made its debut. We have already mentioned the theory of the Davisson-Germer effect and the scattering of alpha particles, both of which properly come under the present item (9).

10. *New light on statistical mechanics.* Because of the fact that the new mechanics is to such a large extent fundamentally statistical in nature, the supplementary assumptions which it is necessary to make to get the fundamental distribution formulas of statistical mechanics are much less violent than in the classical theory. The quantum statistics, of course, merge into those of classical theory at high temperatures, but differ from the latter at low temperatures. The proper quantum statistics to be used in treating radiation are those of the so-called Einstein-Bose type, as this furnishes the Einstein radiation law. On the other hand in dealing with matter it appears necessary to use the Fermi statistics, founded on the Pauli exclusion principle to be discussed later. A particularly significant application is that to electron conduction in solids. If there are about as many conduction electrons as atoms, as seem to be demanded by the optical properties of metals, the classical theory of equipartition would demand that the conduction electrons make an appreciable contribution to the specific heat. The latter would, in fact, be 9 rather than 6 calories per gram atom, contrary to Dulong and Petit's law. In the new quantum statistics, on the other hand, we can suppose that there are as many conduction electrons as atoms without contradicting specific heat data. The reason is that the conduction electrons in solids are relatively close together as compared to molecules in gases, and furthermore the electronic mass is very small in comparison to that of molecules. This means that even at ordinary temperatures the "electron gas" in the solid shows a characteristic quantum "gas degeneration," whereby the specific heat sinks to zero. In a real molecular gas, on the other hand, the degeneration effects would

be important only very near the absolute zero, and even then only at extremely high pressures. Thus the conduction electrons test an aspect of the statistics which it would be difficult to verify otherwise. Pauli shows that the degeneration effects make the magnetism of solid alkalis very small, in agreement with experiment, even though the internal spin makes each electron a tiny magnet and so would give strong paramagnetism with classical theory. Extending Pauli's ideas, Sommerfeld and others have shown that the formulas of thermionic emission, the Wiedemann-Franz ratio, Hall effect, and other properties of conduction electrons are obtainable with the new Fermi statistics. The new vista into electron conduction is, in fact, one of the most pleasing recent developments in quantum mechanics.

IV. WHAT THE QUANTUM MECHANICS PROMISES TO DO FOR THE CHEMIST

Before passing to the chemical applications we must first describe one aspect of the quantum mechanics which seems to be of particular importance for chemistry. This is the so-called *Pauli exclusion principle*, which states that *no two electrons can have simultaneously all four quantum numbers the same*. The reason that there are four quantum numbers per electron is that the electron seems to be a spinning body having a fourth or internal degree of freedom in addition to the three translational degrees of freedom. Just what we take as the four quantum numbers depends somewhat on the type of quantization, which itself varies with the relative magnitude of the various atomic forces. The simplest thing is to suppose that the magnetic field is so powerful that it completely overpowers the inter-electronic forces, so that the orbital and spin angular momentum vectors of each electron are quantized separately relative to the axis of the field. Actually no ordinary magnetic field is powerful enough to do this, but our supposition nevertheless involves no loss of generality because the exclusion principle is a purely formal rule for eliminating certain combinations of quantum numbers, and in any field strength the states will be excluded which pass gradually ("adiabatically") over into states not allowed in ex-

tremely strong fields. In a strong field the four quantum numbers of each electron are the principal quantum number, n , the azimuthal quantum number, k , and the two "magnetic" or "equatorial" quantum numbers, m_k and m_s , quantizing the components of orbital and spin angular momentum in the direction of the applied field. The range of values for the quantum numbers k , m_k , m_s are

$$\begin{aligned} k &= 0, \quad \dots, \quad n-1 \\ m_k &= -k, -(k-1), \quad \dots, \quad 0, \quad \dots, \quad k-1, k \\ m_s &= -\frac{1}{2}, +\frac{1}{2}. \end{aligned} \tag{7}$$

It is to be noted that the azimuthal quantum number is usually numbered one unit lower in the new than in the old quantum theory, as formerly we had $k = 1, 2, \dots, n$. We now have $k = 0, 1, 2, 3$ for optical s, p, d, f terms respectively instead of $1, 2, 3, 4$. (This is purely a formal difference in notation and to emphasize the distinction the notation l instead of k is sometimes used in the new theory). The range of values for m_k is $-k$ to $+k$ because m_k being a component of k , cannot exceed k in absolute magnitude. The quantum number m_s has only the two values $-\frac{1}{2}$ and $+\frac{1}{2}$, because according to the hypothesis of Uhlenbeck and Goudsmit, the internal spin of the electron is associated with a half quantum of angular momentum. Now for given k , equation (7) shows that m_k has $2k + 1$ possible values, while m_s has just two possibilities. There are thus $2(2k + 1)$ different pairs of values for m_k and m_s . Hence one way of stating the exclusion principle is that *in any atom there are at most $2(2k + 1)$ electrons having the same values of n and k .*

Pauli advanced his exclusion principle on more or less empirical grounds before advent of the new mechanics, but the latter has placed this principle on a better and more general basis, as it amounts to using only solutions of Schroedinger's wave equation which are of a certain peculiar type of symmetry (viz., the so-called antisymmetric solutions which change sign when electrons are interchanged). We will not try to enter into the details of the mathematical description of what this type of symmetry is.

The only point that need be emphasized is that *because of Pauli's principle we must not only limit ourselves to "civilized" solutions of Schroedinger's equation, but also in particular only those which are of the "antisymmetric variety."*

Pauli's principle should appeal particularly to chemists, because it yields immediately the well-known interpretation of the Mendeléeff periodic table proposed by Smith and Stoner. These two men showed independently that by making certain apparently rather arbitrary assumptions about the assignment of quantum numbers to electrons of the various groups, one could account very nicely for the structure of the periodic table of the elements—why, for instance, there are 2 elements in the first period, 8 in each of the two next, 18 in the fourth and fifth, but 32 in the sixth. Also within the period many of the chemical properties

TABLE 3

	K-SHELL ($n = 1$)		L-SHELL ($n = 2$)							
n	1	1	2	2	2	2	2	2	2	2
k	0	0	0	0	1	1	1	1	1	1
m_k	0	0	0	0	-1	-1	0	0	1	1
m_s	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$

of the elements, the existence of sub-periods, rare earths, etc., come out nicely. Now Pauli's principle gives automatically the Smith-Stoner assignment of quantum numbers. We will illustrate this, for simplicity, only for the first two periods, i.e., for the K and L shells. In the K shell the principal quantum number n has the value 1, and hence by (7) the azimuthal number k can only have the value zero. On the other hand in the L shell the principal quantum number n equals 2, and by (7) there are the two possibilities $k = 0$ and $k = 1$. The various possibilities for the four electronic quantum numbers are then by (7) those indicated in table 3. The essence of Pauli's principle is that no two columns can have four identical entries. As each column corresponds to a different electron, there are indeed two electrons in the K -shell and eight in the L .

The quantum mechanics should go further than furnish a model of the periodic table. It should tell us what atoms can

combine and what do not. There is no essential difference between the mathematical procedure in the quantum mechanics of calculating a heat of dissociation and of calculating a spectroscopic frequency or critical potential. This fact does not seem quite as universally recognized as it should be. In either case what is required is a knowledge of the magnitude of the various allowed energy levels. The only distinction is that chemical problems involve more than one nucleus, and this considerably increases the labor of calculation. For the latter reason it may well be a long time before the quantum mechanics achieves as many quantitative results for the chemist as for the physicist, but from a standpoint of pure logic there seems to be no apparent reason why the quantum postulates as they now stand should not be adequate to explain the phenomena of chemical reactions. A dynamics which works for the physicist must also work for the chemist and vice versa. It seems scarcely conceivable that a theory which has been so successful in explaining atomic energy levels should fail in the closely related realm of molecular energies. *The mathematical problem of a chemical reaction seems to be this: to investigate whether there are stable solutions of the Schroedinger wave equation corresponding to the interaction between two (or more) atoms, using only the wave functions which have the type of symmetry compatible with Pauli's exclusion principle.*

A beginning in this program of investigating chemical reactions by means of quantum mechanics and Pauli's exclusion principle has been made within the past few months in important papers by London and by Heitler. Although this work is very new, it is already yielding one of the best and most promising theories of valency. The general trend of the work seems to be that because of the critical examination of symmetry properties required by the Pauli exclusion principle, the theory of the classification of valences in complicated organic compounds, etc., must be closely related to the group theory of the mathematicians. Some of the specific results which have already been obtained in the papers of London and Heitler are the following:

1. Two hydrogen atoms can combine to form a hydrogen molecule.
2. Two helium atoms cannot combine, unless excited, without

violating Pauli's exclusion principle. This, of course, agrees with the monatomic property of helium gas. Spectroscopists, to be sure, have observed the spectra due to helium molecules, but such molecules are transient, unstable creatures in which the electrons are in excited rather than normal states.

3. Inert gases cannot exhibit valences.

4. Halides may have the valences 1, 3, 5, 7 except that the valence of fluorine may only have the value 1.

5. S, Se, T may have valences 0, 2, 4, 6 but oxygen only the values 0, 2.

6. P, As, Sb, Bi may have 1, 3, 5 but N only 1, 3.

7. C, Si, Ge, etc., may have 0, 2, 4.

Chemists will immediately recognize that these predicted possible valences are in the main in excellent accord with experiment. Oxygen, for instance, is found to be pronouncedly only divalent, whereas S, Se, etc., have sometimes higher valences. Thus only OCl_2 is observed, whereas on the other hand one finds compounds such as SCl_4 , SF_6 , SeCl_4 , SeF_6 , etc. Similarly it is well known that Cl and I can exhibit higher valences than the unit valence characteristic of F. Some of the valences listed above are not realized experimentally. A notable example is that the valence of nitrogen is never unity, even though according to item 6 the values 1 and 3 are both possible. However, this state of affairs is not a serious difficulty, for I have attempted to list the valences which are possible from the standpoint of elementary symmetry considerations³ rather than those which

³ *Note added in proof.* It is perhaps a bit misleading to state that all of London's results follow unambiguously from symmetry considerations. In order for his definition of valence to be a useful one, it is necessary to suppose a bit empirically that the stable state of a molecule is in most cases one in which the electron spins have zero resultant. The fact that most molecules are diamagnetic shows that this supposition is usually true, but the paramagnetism of the oxygen molecule despite an even number of electrons shows that it does not have universal validity. Calculations with normal hydrogen and helium support London's supposition, but the writer is informed by Prof. Kemble that computations show that in excited states of the hydrogen molecule there is sometimes greater stability with the electron spins uncompensated than with them mutually compensated. Hence London's assumption concerning the predominance of closed spin configurations might well have more justification of a purely theoretical rather

actually exist. I have been very careful to say in every case "may have" according to the theory, rather than "should have." The different valences correspond to different apportionments of various values of the quantum numbers k , m_k among the electrons, and the relative prevalence of the different valences depends upon the relative prevalence of the states corresponding to different values of the quantum numbers k , m_k but given n . Some of these states may have such high energies that they are occupied only very infrequently, and so the corresponding valences may not exist. In other words the preceding inventory indicates only how valency is restricted by the Pauli exclusion principle, and energy considerations may give further limitations. Definite predictions in this respect should be possible if there were available complete spectroscopic evidence on the atomic stationary states corresponding to all assignments of quantum numbers. London's work, in fact, seems to show that there is a very intimate connection between valences and the spectroscopists' classification of spectral terms. Thus he finds that valences 1, 3, 5 in the nitrogen group are associated respectively with doublet, quartet, and sextet structures in the spectroscopists' "multiplets." The spectroscopists find experimentally that the normal state of the nitrogen atom is a quartet rather than doublet spectral term, and this is in agreement with the fact that the normal valence of N is 3. Complete absence of univalence in nitrogen must mean that the lowest doublet terms have considerably greater energy than the lowest quartet terms. Similarly the spectroscopists' observation that the normal spectral term of the O atom is of the triplet rather than singlet type means that normally oxygen is divalent rather than non-valent. This, of course, accords with the fact oxygen gas is diatomic rather than monatomic. Apparently the spectroscopists classifications into multiplets is going to be useful in correlating the relative prevalence of different possible valences.

than semi-empirical nature. It may be noted that Pauling shows (Proc. Nat. Acad. Sci. 14, 359, 1928) that in simple cases London's theory gives results similar to those of Lewis's shared electron pair.

It must also be mentioned that sometimes valences are observed which are not allowed according to the above scheme. Thus nitrogen is sometimes observed to be quintavalent. Such exceptions London explains on the ground that his theory is only one of what he calls "homopolar valences," in which the molecule dissociates into neutral atoms rather than into ions. The valence 5 for nitrogen he claims belongs to the "polar" category in which the molecule breaks up into ions, in which case the symmetry arguments may lead to quite different results. It must be mentioned that London's use of the terms "homopolar" and "polar" may prove a bit confusing, as it has nothing whatsoever to do with the presence of an electric moment, but rather refers to the mode of dissociation. As mentioned in Dr. Darrow's paper in this symposium, recent experimental work shows that molecules are much more apt than has been previously supposed to dissociate into neutral atoms rather than ions. Thus optical dissociation of HBr furnishes H and Br rather than H^+ and Br^- . Hence London would presumably classify the formation of HBr as a homopolar bond despite the electric polarity of the HBr molecule. He is forced to admit that there is no hard and fast distinction between the two terms homopolar and polar as he uses them, since some molecules may dissociate in either of two ways. According to London a valence is of his polar type only if the electron affinity of one atom is greater than the ionization potential of the other.

A point which is particularly to be emphasized is that according to London the reason certain valences or bonds do not occur (e.g., compounds involving inert gases) is not that such bonds lead to molecules which are energetically unstable, but that the bonds, when stable, correspond to solutions of the Schroedinger wave equation which are of a type of symmetry contrary to Pauli's exclusion principle. Thus considerations of symmetry (group theory) are often quite as vital as those of energetics. The failure of a chemist to find a compound does not necessarily mean that the corresponding molecule is energetically unstable, but may mean rather that it would demand electronic groupings contrary to the exclusion principle. As an example consider the

question of whether two normal helium atoms can combine to form a helium molecule. Here the mathematical analysis of Heitler and London shows that there are energetically stable solutions of the wave equation corresponding to the interaction of two helium atoms, as well as also energetically unstable ones. In the stable solutions the statistical charge distribution is roughly of the general type shown in figure 2, whereas in the unstable ones it is more of the type illustrated in figure 3. Hence helium gas might be diatomic, were it not for the fact that the stable solutions are all contrary to Pauli's exclusion principle, leaving only some of the unstable solutions corresponding to repulsion of the atoms rather than chemical affinity. On the other hand in the case of the hydrogen molecule the stable solution satisfies Pauli's principle.

In the application of the exclusion principle to chemical reactions, the existence of an electron spin is of vital consequence. Without the internal spin, for instance, it would be impossible to construct a stable molecule of the hydrogen molecule without violating Pauli's principle. Incidentally, in the hydrogen molecule, the spin axes of the two electrons set themselves antiparallel, so that their magnetic moments compensate each other and make hydrogen gas diamagnetic. Thus indirectly the internal spin of the electron and its attendant magnetic moment is of considerable importance in chemistry. However, one must be very careful not to form the impression that the magnetic forces due to the internal spins are of any magnitude appreciable enough to affect the energy of a chemical bond in the slightest degree. Instead the magnetic forces are entirely negligible in comparison to the electrostatic ones. In the hydrogen molecule, for instance, the magnetic forces are less than 0.1 per cent of the electrostatic. Hence chemical bonds are fundamentally electrical in nature. In this respect the modern Uhlenbeck-Goudsmit spin electron is to be contrasted with the old Parson magneton which pictured the electrons as magnets exerting strong magnetic forces on each other. Instead the internal spin and magnetic moment are important only indirectly in altering the symmetry so that solutions are compatible with Pauli's principle which would not be other-

wise. In other words with an extra degree of freedom and quantum number there are more different ways in which electrons can be arranged without making all quantum numbers the same. It must be confessed that Pauli's dogmatic exclusion principle implying a restriction to antisymmetric wave functions appears decidedly artificial, but is nevertheless the best that can be done in reducing observed spectroscopic and chemical phenomena to a minimum number of postulates, and any ultimate comprehensive theory of matter must presumably include the exclusion principle as one of its consequences.

The London and Heitler theory of valences is primarily based on symmetry properties of the wave functions and does not aim to say anything about the stability or heats of reaction of the various compounds, as this would require a detailed dynamical investigation. There is, however, one chemical reaction in which it has already proved possible to carry through a real mathematical analysis of the dynamical problem on the basis of quantum mechanics. This reaction is the combination of two hydrogen atoms to form a hydrogen molecule. Here only four bodies are involved (two electrons and two protons) so that the equations may really be integrated by rather laborious methods of successive approximations. This has been done independently by Condon, Sugiura, and Wang. The calculated value of the heat of dissociation is 87,000 calories per molecule, whereas the experimental value is 100,100 (spectroscopic) or 97,000 (calorimetric). At first sight this may appear like mediocre agreement, but it must be remembered that the mathematics can be handled only by successive approximations, and a heat of dissociation is a differential effect, viz., the difference in energy between a hydrogen molecule and that of two free hydrogen atoms. These two quantities are nearly equal, so that it is necessary to calculate the energy of the hydrogen molecule very accurately to get the heat of dissociation at all. It is perhaps more expressive of what the calculations have achieved to state the results in terms of the number of volts of energy required to completely disrupt the hydrogen molecule into four separate bodies; namely, two free electrons and two free protons. The calculated value is 30.9

volts and the experimental one is 31.5. The difference of less than a volt is within the error of the calculation. We have here an example of the computation of a chemical affinity by unambiguous methods directly from the quantum postulates, without any undetermined constants or other arbitrary features. Is it too optimistic to hazard the opinion that this is perhaps the beginnings of a science of "mathematical chemistry" in which chemical heats of reaction are calculated by quantum mechanics just as are the spectroscopic frequencies of the physicist? Of course the mathematics will be laborious and involved, and the results always successive approximations. The theoretical computer of molecular energy levels must have a technique comparable with that of a mathematical astronomer. The quantum mechanics is still very young, and surely it will ultimately be applied further than the hydrogen molecule.

SELECTED REFERENCES

General

Schroedinger, *Physical Review*, **28**, 1049-1070 (1926) A very ably written summary by Schroedinger himself of his original German papers.

Born, "Problems of Atomic Dynamics," Chaps X-XX (Massachusetts Institute of Technology, 1926). Written in the early days of the quantum mechanics, but still one of the best accounts of the matrix formulation of the theory, exclusive of its correlation with the Schroedinger wave function. More detail is given in the original German articles, *Zeitschrift fur Physik*, **34**, 858 and **35**, 557 (1925-6)

Haas, "Materiewellen und Quantenmechanik," (Akademische Verlagsgesellschaft, Leipzig, 1928).

Birtwistle, "The New Quantum Mechanics" (Cambridge University Press, 1928).

Biggs, "Wave Mechanics," (Oxford University Press, 1927)

Jordan, *die Naturwissenschaften*, **15**, 614 and 636 (1927). An excellent survey of the different formulations of the theory.

Dirac's monumental, but highly mathematical, papers are found in *Proceedings of the Royal Society*, vol. **109**-date

Heisenberg's Indeterminism Principle

Heisenberg, *Zeitschrift fur Physik*, **43**, 172 (1927).

Bohr, *Nature*, **121**, 580-590 (1928).

The Wave Mechanics

Schroedinger, "Abhandlungen zur Wellenmechanik" (Barth, Leipzig, 1927). Reprint of papers originally in the *Annalen der Physik und die Naturwissen-*

schaften. An English translation has just been published (Blackie and Son, London)

Darrow, Bell System Technical Journal, **6**, 653 (1927). An elementary summary, especially good on the analog to classical vibration problems.

Application to Physical Problems

A bibliography on this would be too extensive to be feasible. The reader is referred to two reports by the writer in Journal of the Optical Society of America, **14**, 108-112 (1927) and **16**, 301-306 (1928). These are portions of the report of the progress committee of the Optical Society and are only brief summaries of the main physical applications during the preceding year. They are, however, documented fairly completely with the appropriate references to the literature.

Chemical Application

Pauli's original paper on his exclusion principle is Zeitschrift für Physik, **31**, 765 (1925) Summary in Hund, "Linienpektren," pp. 114-118.

Hertler and London, Zeitschrift für Physik, **44**, 455 (1927). Impossibility of forming normal helium molecule, etc.

London, Zeitschrift für Physik, **46**, 455; **50**, 24 (1928). Valency rules

Heitler, Zeitschrift für Physik, **47**, 835 (1928). Group theory in relation to chemical bonds.

Wang, Physical Review, **31**, 579 (1928) and references to earlier work by Condon and Sugiura. Heat of dissociation and other properties of the hydrogen molecule.

During the printing of this paper, an excellent summary of the application of the quantum mechanics to the hydrogen molecule, molecule-ion and related problems has been given by L. Pauling, Chemical Reviews, **5**, 173 (1928).

THE ELECTRON THEORY OF VALENCE

WORTH H. RODEBUSH

Department of Chemistry, University of Illinois, Urbana

The scientist has of late been repeatedly warned by the philosophers that he should be careful about pushing the limits of his domain too far, lest he find himself in the forbidden territory of metaphysics. Those of us who are interested in the nature of matter may be startled to find that in our speculations concerning the atom we have raised questions of a philosophical nature, which we shall probably not be able to answer, without going outside what is considered by many to be the proper field of science. This is perhaps only a dignified way of saying that they can not be answered at all.

Bridgman (1) points out that the atom is a "construct;" that is, a mechanism designed to explain the behavior of matter. The atom seems destined forever to escape direct observation; the evidence for its existence is wholly inferential. The most important atomic property that is directly measured is electric charge; as it is defined the measurements indicate that it exists in discrete units. The ratio of e/m is of course a result of measurement; but the mass appears to vary with speed because the charge is assumed to be constant. The proton and electron have come to seem very real to us but actually they are "constructs," which represent a long extrapolation beyond the atom, which itself cannot be seen. The most difficult problem of the physicist is to account for the structure of the atom itself. The only experimental law, which he has for guidance, is Coulomb's law of the force varying as the inverse square of the distance between charges. This law is wholly insufficient to account for the structure of the atom or even the stability of the electron and proton. Why the electron does not explode spontaneously, or fall into the proton with catastrophic results at the first approach is not

explained. Eddington (2) and others have speculated on the possibility of this happening. Any attempt to explain the fundamental behavior of electrical charges must involve the assumption of a modification of the Coulomb law of force; such a hypothesis has been elaborated mathematically by Birkhoff (3). The Bohr atom, as was pointed out by G. N. Lewis (4), in addition to assuming the proton and electron of the Rutherford atom, assumes lapses at certain times and places of the laws of electrostatics and electrodynamics. That the proton and the electron are condensed into infinitesimal volumes is itself an assumption. The Schroedinger wave equation has been interpreted to indicate a diffuse electron but it seems more satisfactory to interpret it as a sort of smoothed out or continuous function which indicates the discontinuous probability of location in time and space of the corpuscles of matter and radiation. Since the atom is a "construct" or hypothetical mechanism invented to explain the behavior of matter, it should be designed in the simplest possible way that will serve the purpose. The Bohr atom was designed to explain radiation only and was never a satisfactory model for the chemist. It seems more than a coincidence that Born, the leader of the revolution in thought known as the new quantum theory, is a man who has a large familiarity with the facts of chemistry. The statement that the orbit exists only when it can be observed sounds like a paraphrase of the statement made by G. N. Lewis when he said more than ten years ago that an electron should not be said to be in motion when it had none of the properties of motion. Heisenberg's (5) Principle of Indetermination in which he denies the possibility of locating an electron at a given instant is somewhat reminiscent of Ostwald's denial of the existence of the atom itself. Ostwald led the physical chemists on what appears now to have been an inglorious retreat. The refusal to make pictures, however much it may seem to be justified on philosophical grounds, or those of common sense, appears to be a sterile procedure from the standpoint of advancing knowledge. The invention of models, no matter how crude and artificial they may be, seems justified by the experiments that are inevitably suggested. The results of these ex-

periments seldom confirm the hypothesis, or worse still they may not even answer the question in an unambiguous manner, but they almost always give us new and interesting information.

The chemist and especially the organic chemist has accumulated the largest amount of consistent data without a rational underlying theory that a science has ever known. The physicist has constructed the most intricate theory of atomic structure, largely from the data of spectroscopy, that the history of science records. Until recently the theory of the physicist and facts of the chemist have seemed a long way apart. Within the last few years the two have been drawing together very rapidly and in this process the physicist seems to have yielded more ground than the chemist. The facts of chemistry seem more obvious and more direct in inference than the data of spectroscopy. The physicist appears to have learned more from the chemist than the chemist has learned from the physicist. The physicist now tells the chemist that his ways of looking at things are really quite right because the new theories of the atom justify that interpretation, but of course the chemist has known all the time that his theories had at least the justification of correspondence with a great number and variety of experimental facts.

The model of the atom currently accepted by the physicist is a Rutherford nucleus about which the electrons are arranged in groups or shells, of substantially the same energies, which correspond to the periods 2, 8, 8, 18, etc., of the Thomsen arrangement of the periodic table. The outermost group never contains more than eight electrons. With the underlying group the chemist is never directly concerned for no fact seems more certain than that the valence electrons are in the outermost group. These shells are commonly designated, started from the nucleus outward by the letters *K*, *L*, *M*, etc. To the physicist they represent merely a classification of the electrons into levels of approximately the same energy and are commonly designated by the principal quantum numbers 1, 2, 3, etc. According to the Main Smith-Stoner scheme of electron arrangement the completed outer group of eight electrons in an inert gas may be divided into three sublevels which are filled by 2, 2 and 4 electrons respectively.

This arrangement agrees with the spectroscopic evidence; the agreement with the chemical evidence will be discussed later.

With such structure postulated for the atom and Coulomb's law one can do much toward explaining the formation of molecules in a qualitative way. G. N. Lewis (4) has stressed the complete failure of the electrochemical theory of Berzelius in accounting for the formation of any considerable number of compounds. It must be remembered, however, that the electrochemical theory was invented at a time when the atom was supposed to be made up of matter and electrical charge was supposed to be acquired more or less incidentally, as a dust particle might adhere to a larger object. When it is recognized that the atom is made up primarily of electrical charges, then the theory must of course be greatly modified and if so modified, the theory is by no means so inadequate as has been stated. The mutual attractions between the nucleus of one atom and the electrons of another are quite sufficient to account for the energy of formation of molecules, provided a suitable arrangement of the electrons around the nucleus is assumed. Thus an excuse is to be made for the failure of the physicist to calculate the energy formation of sodium chloride from the ionizing potential of sodium and the electron affinity of chlorine. These quantities are measured for the respective atoms in the gaseous state, but sodium chloride in the gaseous state is not a polar compound in the sense that the electron is completely transferred from the bond between the sodium and the chlorine atoms. It must be noted to the credit of the physicist that he is now able to calculate heats of dissociation of diatomic molecules¹ and that the Schroedinger wave equation has been used to calculate the energy of formation of the hydrogen molecule. The outstanding difficulty then in a theory of valence is not to account for the forces which bind the atoms into molecules. The electrochemical theory can do that. It is rather to account for the existence of some compounds, the nonexistence of others and particularly, the discrete unitary

¹ Mr E. C. Walters working with the author at the University of Illinois has obtained results confirming the prediction of the physicist that sodium vapor contains diatomic molecules.

nature of valence, which can be expressed by a series of small whole numbers for the chemical elements and consequently, leads to the law of multiple proportions. These characteristics of valence are accounted for by the brilliant theories of G. N. Lewis (4) in a remarkably satisfactory manner, at least from the chemist's point of view. Lewis has formulated his theory in part as a set of numerical rules, concerning the disposition of the valence electrons in the atom and molecule. These rules must appear, however, in a sense, arbitrary, until they are related to the structure of the atom as the physicist pictures it. We shall state these rules in turn and consider the theoretical basis for each.

THE RULE OF TWO

Lewis was led to his conclusion that the electrons tend to arrange themselves in pairs by the extraordinary significance which even numbers appear to have for the arrangement of the electrons both within and without the nucleus. The following are some of the generalizations.

1. Even number of electrons in the nucleus (usually).
2. β -decompositions of radioactive elements occur in pairs.
3. Very few "odd" molecules.
4. Ionizing potential of odd-numbered elements usually lower than adjacent even-numbered elements.
5. All odd-numbered elements are paramagnetic.

The most important of these generalizations for the chemist is undoubtedly the third one; namely the scarcity of molecules containing an odd number of electrons. From this Lewis was led to the theory that the nonpolar valence bond is a pair of electrons occupying a position intermediate between the two atoms. This theory does not follow from the generalizations above by direct inference any more than the atomic theory can be inferred from the laws of chemical combinations but it appears to be the only explanation of these facts that any one can offer. Since according to Coulomb's law two electrons should exert a repulsion for each other, the pairing of electrons seems at first glance

to be a bizarre idea indeed. In order to account for the peculiar behavior Lewis assumed the existence of a magnetic attraction between the electrons. One of the most striking of the experimental generalizations is (5) the fact that all atoms of elements of odd atomic number are paramagnetic while the atoms of most even numbered elements are not. Lewis has attributed a good deal of importance to the generalization that the ionizing poten-

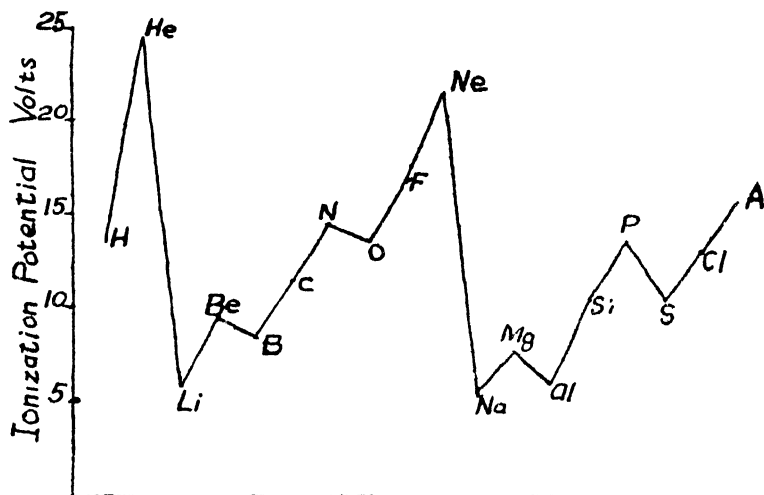


FIG. 1

tials of the odd numbered elements are usually higher than the adjacent even numbered elements. Figure 1 is a curve showing the ionizing potentials of the first three periods of the elements according to Millikan and Bowen.² It must not be assumed however that magnetism is a simple, readily comprehended explanation of the pairing of electrons. Magnetism in the atom is intimately connected with the peculiar phenomenon of space quanti-

² Millikan and Bowen, *Proc Nat Acad. Sci* **13**, 531 (1927) It will be noted that nitrogen and oxygen are reversed in the curve; likewise, phosphorus and sulfur. Nitrogen contains an odd number of electrons in the nucleus, while oxygen is anomalous in its magnetic properties. Unfortunately, the explanation is not applicable to phosphorus and sulfur.

zation. Instead of a random orientation with respect to the field simple paramagnetic atoms are oriented in parallel and antiparallel positions to the field. The parallel position is of course plausible enough but the antiparallel position seems more mysterious. A little reflection will show, however, that if the mass susceptibility is to remain constant that this partition between parallel and antiparallel must approach equality as the field is decreased to zero. This orientation gives rise to the peculiar results of the Stern-Gerlach experiment in which a beam of atoms passing through an inhomogeneous magnetic field is split into two components. Results have been obtained in this laboratory^a (6) recently on atomic hydrogen, sodium and potassium, which show the typical behavior of paramagnetic atoms. A further difficulty in explaining this experiment is in accounting for the energy necessary to produce the orientation since the atom requires only 10^{-4} seconds to traverse the field. This energy is more easily accounted for if instead of the atom being oriented as a whole the electron itself possessed a magnetic moment and is oriented independently of the atom.

The magneton or magnetic electron has been invented repeatedly to account for atomic phenomena although it has not yet been discovered, i.e., its existence demonstrated by direct experiment. Parson (7) was one of the early writers to suggest the significance of the magneton in chemical valence. Recently Uhlenbeck and Goudsmit (8) have suggested that the electron spins on its own axis with a velocity sufficient to give it a magnetic moment $\left(\frac{eh}{4 \pi mc} \right)$ equal to the Bohr magneton. The Bohr magneton is of course the unit magnetic moment which would be associated with an electron moving in an orbit with one quantum of angular momentum.

The spinning electron appears to have been an extraordinarily happy suggestion to both the physicist and the chemist. It furnishes a basis for Pauli's Exclusion Principle (9). The electron was formerly presumed to have three degrees of freedom and was

^a Laboratory of Physical Chemistry, University of Illinois.

assigned three quantum numbers, n , the "principle" quantum number locates the electron in its shell or group; 1 gives the number of quanta of angular momenta associated with the motion of the electron and m_l gives the component of angular momentum in the direction of an external magnetic field. Pauli assumes another degree of freedom connected with the spacial orientation of the spin axis and assigns a fourth quantum number $m_s = \pm \frac{1}{2}$, designating parallel and antiparallel orientations. Pauli's Principle states that no two electrons in an atom can have identical values for all four quantum numbers. Stated in another way this means that there will be a pair of electrons corresponding to identical values of the first three quantum numbers in the atom, but these two electrons will differ in that their spin axes are oriented in opposite directions to each other. In other words they will neutralize each other magnetically. Hence every atom with an odd number of electrons must contain at least one free magneton. Taken in connection with the possible values of the quantum numbers for the different shells,—Pauli's Exclusion Principle limits the total number of electrons possible in the atom and gives the arrangement of 2, 8, 8, 18, etc., together with the Main Smith-Stoner scheme of sublevels. It is an extraordinarily simple and comprehensive rule and in that sense may be said to be logically satisfying, in so far as any principle connected with the quantum theory can be logically satisfying.

It will be recognized by the chemist however that Pauli's rule is only a short hand way of saying what Lewis has assumed for many years as the basis of his magnetochemical theory (10) of valence. If the electrons are paired in the atom magnetically, it is easy to see how two unpaired electrons in different atoms may be coupled magnetically and form the nonpolar bond. The exclusion principle is therefore extended to molecules and Pauli and Somerfeld have even proposed to extend it to the conducting electrons in a whole crystal lattice of a metal. The chemist has been inclined to attribute this energy of formation of the electron pair to the magnetism of the electrons. The physicist says that this is not true; that the magnetic energy is so small as to be quite negligible. London (11) in a recent paper has

shown that the energy associated with a pair of electrons, magnetically coupled between two atoms is due to a "resonance" effect discovered by Heisenberg (12). It results from what may be pictured as an oscillatory interchange of the electrons between the two atoms so that the electron appears to be a part of both atoms. This evidently corresponds to what the chemist means when he says a pair of electrons is shared between two atoms. This "exchange energy" is negative and binds the two atoms together if the electrons have their spin axes opposed so as to neutralize each other magnetically, but if they do not neutralize each other magnetically it leads to a repulsion between the two atoms. The energy of the molecule thus depends upon the magnetic pairing of the electrons and in dealing with anything so elusive as energy in a quantized system, it appears to the layman somewhat rhetorical to say that the energy depends upon the magnetism but is not magnetic in origin. At any rate the success of the paired electron theory in assigning valence structure to the hundreds of thousands of compounds in chemistry is so well known to chemists as to need no further recital here.

THE RULE OF EIGHT

The rule of eight appears to be almost as universal in its validity and as important in its consequences as the rule of two. Strictly speaking it should be stated as follows: every atom in a chemical compound tends to take on a configuration of its valence electrons, similar to that of the nearest inert gas. This is two for helium and eight for neon, argon, etc. The change from two to eight takes place somewhere in the middle of the second period of elements and it is possible that carbon and nitrogen sometimes form stable groups of six electrons. There are probably not many more exceptions to the rule of eight than to the rule of two. On the other hand the physical basis for the rule of eight seems to be more obscure than for Pauli's Exclusion Principle. However mysterious the Pauli Exclusion Principle may seem, it has the obvious physical basis of the pairing of magnetons. From the rule that the auxiliary quantum number 1 may not be greater than the principle quantum number decreased by one ($1 \leq$

$n-1$), Pauli's Exclusion Principle limits the possible number of electrons for helium to two and the possible number for neon to eight. For argon and the other inert gases however no such limitation appears to exist, and we might expect that more than eight electrons would be found in the outermost group. The assignment of a definite number of energy levels and sublevels which are designated by the various combinations of quantum numbers must seem to the chemist a very arbitrary procedure although it is possible to predict these, presumably, by calculations from the Schrodinger wave theory. This has only been done for the simplest molecule, hydrogen, since the mathematical difficulties are very great for the more complicated atoms. The great importance for the rule of eight to the chemist is that it determines

	Li	Be	B	C	N	O	F	Ne
<i>Electrons</i>	1	2	2,1	2,2	2,2,1	2,2,2	2,2,3	2,2,4
<i>Valences</i>	1	2	3,1	4,2	5,3,1	6,4,2	7,5,3	

FIG 2

the possible chemical compounds. The vast majority of the predicted compounds have been made in the laboratory.

The first attempt of the physicists to calculate possible valences for the various elements was made in connection with the Main Smith-Stoner scheme of sublevels already referred to. According to this scheme in passing from an alkali metal at the beginning of a period of eight to the inert gas at the end of the period, electrons are added successively in the subgroups, of two, two and four. The idea proposed was that the lowest valence of an element would involve all of the electrons in the last subgroup; the next higher valence would include the electrons of the second subgroup and the highest valence would include all of the valence electrons. Reference to figure 2 will make this clear. Main Smith (13) claimed that this scheme agreed with the chemical evidence and Grimm and Sommerfeld (14) published a paper in which they apparently found a correlation. The agreement,

however, is far from satisfactory. In the first place the fact that oxygen and fluorine show but one valence is entirely unaccounted for. Furthermore chlorine has a valence of one, nitrogen perhaps never has a valence of one, and boron does not have a valence of one. All three cases disagree with the scheme. On the other hand the rule of two will account in a very satisfactory manner for the behavior of these elements with the exception of the absence of a valence of one for boron. The rule of eight explains nicely the practical nonoccurrence of a valence of one for nitrogen.

Recently London (11) has discussed the valence of these elements in a very interesting paper. He is evidently thinking in terms of the old valence theory, in which an element has one valence toward oxygen and another toward hydrogen, which will prejudice the chemists against his ideas somewhat, since these old valence theories have proved inadequate. Also London talks in terms of heteropolar and homopolar valences, which we shall point out later is not a happy terminology. London makes one assumption which will evidently lead to radically different results from the Lewis theory, namely that, when a pair of electrons is shared between two atoms, one electron must be contributed from each atom. This leads him to the conclusion apparently that the chlorine in perchlorate ion shares seven pairs of electrons with the oxygen, thus doing away with the rule of eight. It is not clear why chlorine should not show a valence of seven towards other elements such as hydrogen if it does so for oxygen. The chemist is inclined to regard the combination of oxygen and chlorine in perchlorate ion as analogous to the combination of hydrogen and nitrogen in the ammonium ion and will be loathe to abandon the rule of eight for this compound. The most interesting part of London's paper is perhaps his attempt to show why nitrogen, oxygen and fluorine do not have as many different valences as phosphorus, sulfur and chlorine. He does this by showing that the Pauli Exclusion Principle taken together with the possible quantum numbers for nitrogen, oxygen and fluorine limits the number of unpaired electrons in these atoms to three for nitrogen, two for oxygen and one for fluorine. This is undoubtedly cor-

rect but if we accept his principle that a pair of electrons, belonging to one atom originally, cannot act as a valence bond then oxygen cannot form a complex analogous to the ammonium ion, whereas the H_3O^+ ion certainly exists in solution. The chemist would prefer to assume that any pair of electrons can act as a bond and explain the nonexistence of higher valences for nitrogen, oxygen and fluorine as being due to energy considerations. The chemist does not require a quantum prohibition to explain the non existence of nitrogen pentachloride knowing as he does the properties of nitrogen trichloride. He would rather have an explanation for the instability of nitrogen trichloride. London's explanation of nitric acid as a polar hydroxide is somewhat far fetched. His explanation of the absence of univalent nitrogen is not obvious but he relates it to the fact that the spectroscopic ground term for nitrogen indicates three unpaired electrons. It is evident that London's reasoning would wipe out all differences in degree of stability between chemical compounds and substitute the hard and fast discontinuities of the quantum theory. There is much justification for this view from what we know about the atom and it may be that the rule of eight is a sort of chemical correspondence principle which seems to work because our methods of measuring chemical properties are essentially qualitative. The fact that the rule of eight works well is not absolute proof of its correctness; the old valence theories could be formulated in rules which had a wide applicability. On the other hand a rule which works for the hundreds of thousands of compounds with so few exceptions as the rule of eight is not to be discarded lightly and the chemist who has observed the variation of properties through a series of similar molecules will be loathe to discard the idea of differences of degree in stability. For the present at least the chemist will be inclined to continue to use the the rule of eight for what it is worth.

THE RULE OF FOUR

The rule of eight leads to the rule of four, that the maximum number of atoms which may be attached to a central bond by nonpolar bonds is four. This is obviously a limiting rule only.

Under the old valence rules the emphasis is upon the number of atoms linked to the central atom. In the new theory an atom is "saturated" when it has eight electrons regardless of the number of atoms to which it is linked. This is one of the most radical differences between the old and the new theories and greatly increases the number of stable molecules possible. This rule makes possible such things as bivalent carbon and univalent oxygen which were absurd under the old theory. This gives us tetrahedral symmetry and the beautiful stereochemistry of Van't Hoff which has been one of the triumphs of organic chemistry. Recently x-ray studies of pentaerythrite have indicated that the carbon atom is at the apex of a pyramid with the four carbinol groups forming the base. Weissenberg (15) and Henri have made much of this, claiming that the stereochemistry is all wrong and even Willstätter seems to have been impressed with the theory. The burden of proof is upon the new theory and the x-ray evidence in the case of pentaerythrite is disputed. The fact that methyl chloride and chloroform have electrical moments, while carbon tetrachloride does not, would seem to be almost convincing evidence for the tetrahedral structure of the carbon tetrachloride molecule.

POLARITY

While it is important to account for the structure of the known chemical compounds by a theory of valence, it is of even greater practical importance to predict the chemical properties and reactivity of molecules once their structure has been assigned. In the old valence theories a bond was a rigid unyielding link, the same whenever it was found. Now that we picture the bond as a pair of electrons held between atoms of different nuclear charges, we begin to see how the properties of the bond must vary from one molecule to another. One of the most important of the factors which determine the character of a bond is the degree of polarity. We may remark parenthetically that the use of the terms homopolar and heteropolar does not seem to make for clearness in discussing the structure of molecules. We have polar compounds such as sodium chloride in the crystal state

where the sodium is a positive ion and the chlorine is a negative ion and no molecules appear to exist. All other molecules appear to be made up of atoms joined together by nonpolar bonds, that is, pairs of electrons.⁴ The terms positive and negative have been used in so many ways in connection with the properties of molecules that they are likewise ambiguous. The words may be used in two connections. In the first place, we may wish to designate the charge on an ion as the chlorine in sodium chloride. We can avoid ambiguity here by saying that the chlorine is negatively charged. The terms positive and negative may also be used to characterize the field surrounding the atom. At a distance the field of a neutral atom is zero, but the chemical properties depend upon the field in the immediate neighborhood of an atom. With positive and negative charges grouped together in an atom but separated from each other by finite distances we obtain complex fields which vary as the inverse cube or higher powers of the distance from the atom.⁵ The field in the neighborhood of all atoms except the inert gases is positive, that is, the atom will attract electrons. The higher the effective nuclear charge⁶ the more positive the atom. Thus sodium would be designated as weakly positive and chlorine as strongly positive. This is contrary to the current usage in inorganic chemistry where chlorine is usually termed negative or electronegative, but the current usage arose before the discovery of the electron and there seems to be no good reason for adhering to an illogical terminology.

The term electroaffinity was used many years ago to characterize this important property of an atom, i.e., its field at small distances, and there seems to be no better term available for the purpose.

The ionizing potential of an atom appears to be a rough measure of its electroaffinity. Elements whose ionizing potentials are

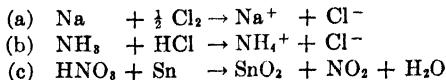
⁴ Thus sodium chloride in the vapor state is apparently a molecule with the sodium ion joined to the chlorine by a pair of electrons. London quotes Franck as having advanced the idea that the hydrogen halides are nonpolar in the pure state, but this is scarcely a new idea to the chemist.

⁵ Thus it can be shown from Coulomb's law that the field due to a dipole varies as the inverse cube of the distance.

⁶ The effective nuclear charge is equal to the atomic number diminished by the screening effect or repulsive action of all the other electrons of the atom.

above ten volts are distinctly non-metallic in character while those whose ionizing potentials are much below are characteristic metals. Hydrogen is thus clearly classified with the halogens rather than with the alkali metals. The electroaffinity of the atoms is the most fundamental factor in determining the chemical properties of compounds, especially in inorganic chemistry. Many of the more characteristic properties of molecules depend upon the combination of two atoms with considerable difference in electroaffinity. If the difference is great enough we may get a "completely polar" substance like sodium chloride where the molecule is split into ions and the analogy to a dipole has disappeared. On the other hand, in a large number of molecules such as water and hydrogen chloride there is only a slight degree of polarity which may be confirmed by the calculation of the electrical moment from the temperature coefficient of the dielectric constant. Too much significance must not be given to the electrical moment in the gaseous state. It is not possible for instance to predict from it the dielectric constant in the liquid state. We can see how this comes about by considering the water molecule. If the water molecule had a symmetrical structure (atoms in a straight line) its electrical moment would be zero. Most of the important properties due to polarity appear only in the liquid state or in solution. Where we may expect a slight degree of "internal" polarity to be greatly augmented by the influence of adjacent molecules.

Atoms or groups of atoms of high electroaffinity are conspicuously oxidizing and acid forming in their behavior. Atoms of low electroaffinity are base-forming and reducing. The relation oxidation and neutralization is shown by the following type reactions.



(a) is a typical oxidation reaction; an electron is transferred from the sodium to the chlorine.

(b) is a typical neutralization; a hydrogen ion is transferred

from the chlorine to the ammonia. But the hydrogen ion is the proton or positive electron and the transfer of positive electricity in one direction is the same as the transfer of negative electricity in the opposite direction. (a) and (b) are therefore completely analogous. (c) appears somewhat different in that a neutral oxygen atom is transferred but if we remember that the oxygen atom is "deficient" in electrons it becomes analogous to the other two reactions; from the standpoint of the transfer of electricity. By deficient in electrons we do not mean that the oxygen is not electrically neutral but rather that it carries but six electrons while it seeks a complete group of eight. When oxygen and nitrogen combine with each other or with carbon there is usually a deficiency in electrons, i.e., there are not enough electrons to give each atom a complete group of eight without sharing more than one pair of electrons between two atoms. Such a group usually behaves as if "unsaturated" and strongly positive. The only elements which can replace oxygen in a reaction of type (c) are those which form double bonds. Nitrogen and sulfur are known to act in this way but carbon apparently does not.

POLARITY IN ORGANIC REACTIONS

When we attempt to predict the relative positivity of two atoms joined by pairs of electrons we must recognize a peculiar effect. The displacement of the electrons from the less positive to the more positive atom or group may more or less completely compensate the difference in positivity. At first thought we might expect the positivity to be exactly compensated, but the quantum theory would lead us to expect the actual displacement would be such as to produce either over or under compensation and the facts of organic chemistry appear to confirm this. In order to decide the relative positivity of neighboring atoms then we shall have to scrutinize each individual case.

It may be remarked here parenthetically that the designation of positive and negative atoms in an organic chain by plus and minus symbols appears to be undesirable since it suggests the positively and negatively charged ions of inorganic chemistry.

We certainly have no degree of polarity in ordinary organic compounds which approaches that of sodium chloride.⁷

Since we cannot detect slight polarities by physical means we are forced to rely on chemical reactions. The most reliable reagent for polarity appears to be the Grignard reagent. The large number of compounds with which the Grignard reagent reacts and the speed of the reaction are characteristics which are suggestive of ionic reactions. Indeed it has been suggested (16) that the compound RMgX is a typically polar compound in both the Mg-R and the Mg-X linkages. It is well known that ether solutions of the Grignard show considerable electrical conductivity (17). In the hope of demonstrating by physical means the extent of the polarity of the bond between a metal and an organic radical, numerous experiments in the electrolysis of compounds in which this bond is present, have been made by Dr. J. M. Peterson working with the author in the Laboratory of Physical Chemistry at the University of Illinois.

Repeated attempts were made to isolate the products of electrolysis from ether solutions of magnesium ethyl iodide and magnesium phenyl iodide. At the cathode magnesium is precipitated as might be expected but the liberation of iodine at the anode masked completely any effects due to the organic radical, although Gaddum (18) has apparently been more successful in similar experiments.

In order to avoid the effects due to the highly ionized halogen, zinc diethyl was substituted for the magnesium compounds in these experiments. Any inference which can be drawn as to the polarity of zinc diethyl must apply in some measure to the Mg-R linkage since the two metals are undoubtedly very similar in their behavior. Hein (19) has shown that sodium ethyl dissolved in zinc diethyl is a conductor but that pure zinc diethyl is not. This latter conclusion was confirmed by us. Zinc diethyl when carefully prepared and purified by distillation in a dry inert atmosphere does not show an appreciable conductivity. If, however, a small amount of pure dry ether is distilled into the cell

⁷ For a discussion of the properties of highly polar compounds see Latimer and Rodebush, *J. Am. Chem. Soc.* **42**, 1419 (1920).

containing the zinc ethyl a specific conductivity between 10^{-5} and 10^{-4} is at once obtained and using 110 volts appreciable quantities of zinc may be deposited on the cathode. The hydrocarbons which must be formed at the anode are of course dissolved in the solution and not easily detected in small amounts. Since the ether has no conductivity the ions which carry the current must come from the zinc diethyl and must be $(\text{ZnC}_2\text{H}_5)^+$ and C_2H_5^- (possibly Zn^{++}).

Undoubtedly sodium ethyl if it could be brought into a liquid condition would show the ionization and conductivity of a typical polar fused salt. Magnesium and zinc however, belong to a group of metals of which the two lower members cadmium and mercury show slight ionization in combination with halogen. Magnesium and zinc show a decided tendency to coördinate auxiliary valence groups as does cobalt in the cobalt amines. Ionization in the salts of these metals appears to take place readily only when the anion is replaced by a molecule of the solvent. The bond between the zinc and ethyl radical is essentially a nonpolar linkage and it is only when the oxygen of the ether is present to offer its electrons as a substitute, that the zinc ion lets go of the ethyl group. Since the dielectric constant of the solution is low the separation of the ions is not favored and the conductivity will be low unless ions are present in great numbers. A much greater ionization could be expected if we could introduce a solvent like ammonia or water without disturbing reactions.

The same conclusions must apply to the alkyl or aryl magnesium halides. The Mg-X bond must possess a considerable degree of polarity. The negatively charged ion R^- where R is a saturated aliphatic hydrocarbon radical is the most basic of any group that we know, using the term basic in the sense of readiness to share electrons. Magnesium has a considerable affinity for electrons and the bond Mg-R is to be regarded as distinctly of the non-polar type. If this bond were strong no dissociation could occur. But the effective nuclear charge of the magnesium atom is only 2^+ and the bond does not possess the rigidity of a carbon to carbon bond. Dissociation may take place especially if the magnesium atom transfers its attraction to the

electrons of the oxygen of the ether molecule. The fact that the Grignard reagent combines with two molecules of ether is evidence that this tendency exists. The amount of the dissociation is a statistical property of the relative readiness with which the organic radical ion and the oxygen of the ether share electrons with the magnesium atom. Since oxygen is less basic than nitrogen, and much less basic than carbon, we should expect the dissociation to be slight. However, the dissociation need not be large in order for a reaction of an ionic character to proceed rapidly. For instance aluminum chloride reacts rapidly with water although neither substance is highly ionized. From the foregoing it might be predicted that the Grignard reagent would be more reactive in ether than in a more inert solvent such as benzene, for the benzene molecule would not attach itself so readily to the magnesium ion. This prediction appears to be borne out by the facts.

The simplest reaction that the Grignard reagent undergoes is the addition of acid hydrogen to the basic hydrocarbon, a reaction that is exactly analogous to neutralization except that the product formed is a hydrocarbon instead of water. Most other reactions can be classified as the addition of the negatively charged hydrocarbon radical to the positive atom of an unsaturated pair, a reaction that is analogous to the hydrolysis of aluminum chloride.

Typical examples of such pairs are $C=O$, $C=N$, $N=O$. Here in each case the Grignard radical adds to the first atom of the pair. The part played by the magnesium and halogen are of no significance since they are essentially ions. It will be noted that in each case the second atom of the group has the higher nuclear charge and should be the more positive but the electron displacement has more than compensated for the difference. Thus the carbon atom in $C=O$ is the more positive atom. It is very striking that while we began by using the term positive to mean exactly the opposite of its common use in inorganic chemistry we end by finding our usage to coincide exactly with the long established usage of organic chemistry.

There appear to be no important exceptions to the above rule

of behavior of the Grignard reagent and it constitutes a convincing proof of an appreciable degree of polarity in organic linkages. On the other hand the Grignard reagent does not add to the $C=C$ bond and the reason here seems to be that this bond is symmetrical and the deficiency of electrons being divided equally between the two atoms is not great enough in either to pull in the R : group and bring about a reaction.

The case of the ethylenic linkage becomes particularly clear when we consider a conjugated system such as butadiene where 1, 4 addition usually takes place.⁸ The reason for this can be readily seen. If we regard the 3, 4 linkage simply as a group deficient in electrons then its effect on the 1, 2 group is to displace the electrons toward the second carbon atom leaving the first carbon atom positive. In a similar way the 1, 2 linkage causes the fourth carbon to become positive by displacement of electrons toward the third. Atoms carrying electrons will then tend to add to the two end carbons. As soon as the addition takes place the excess electrons in the 1-2 and 3-4 bonds will shift to the 2-3 position. Thus the Grignard reagent may be made to react with the ethylenic linkage in some conjugated systems.

The theory of alternating polarities of the atoms in a carbon chain does not appear to be substantiated by the experimental evidence. Instead of assuming that the carbons in the benzene ring have alternately six and eight electrons, it seems much more reasonable to assume a conjugation of ethylenic linkages. That binary groups do not add 1, 4 as in the case of butadiene may be explained by assuming that for the addition of the binary group the 1, 4 atoms must come in juxtaposition and this would be prevented by the rigidity of the six membered ring.

REACTIVITY IN ORGANIC COMPOUNDS

The difficulty of explaining reactivity in organic compounds is illustrated by the following typical case. A halogen attached to the carbon of a double bond is very inert; while a halogen at-

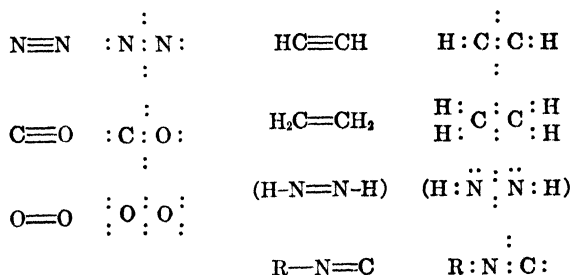
⁸ Lowry (16) writes butadiene in the following manner $CH_2^+ = CH^- - CH^+ = CH_2^-$. If we were to use plus and minus signs, we should prefer $CH_2^+ = CH^- - CH^+ = CH_2^-$.

tached to a carbon adjacent to the double bond shows a high reactivity. The first case is plausible. The halogen can only be removed as a negative ion and the carbon holds on to the electrons. The second case seems at first thought contradictory.

There seems to be no justification for any theory of alternating polarity of the carbon atoms. What we are forced to assume is that under strain from the action of opposing groups (here the halogen and the double bond) oscillations in the electron systems are set up. In the case of the halogen attached to a double bond, these oscillations are not of sufficient amplitude to rupture the bonds. In the allyl chloride on the other hand with a greater inertia the frequency of oscillation corresponds to that of the molecular bombardment and the amplitude of oscillation becomes great enough for the halogen to break loose.

MULTIPLE BONDS

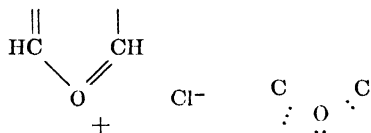
There appears to be a great deal of confusion in the minds of many as to whether a double bond is twice as strong as a single bond or whether Bayer's strain theory holds a multiple bond should behave as if unstable. The question appears to be solved by considering the two classes of multiple bonds.



The first group are characteristically inert and stable while the second group behave as if unsaturated. Evidently two small atoms may combine to form a stable molecule but if a third and fourth atom are attached a strain is produced. The fact that the carbide ion $\text{C}\equiv\text{C}^{--}$ in sodium and calcium carbide is stable

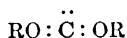
while the less polar silver and copper acetylides are explosive agrees with the above generalization. The structure for carbon monoxide resembles that of the nitrogen molecule; the two atoms are connected by three pairs of electrons. This structure accounts for the saturated behavior of the gas. It might be supposed that the molecule would be unbalanced electrically. The molecule does show a small electrical moment but due to overcompensation which has been mentioned previously the electrons are probably shifted toward the oxygen and the carbon is probably the more positive atom. The cyanide (isocyanide) ion should have a structure similar to carbon monoxide.

While not the most important, one of the most interesting applications of the electron theory of valence is in explaining the structure of compounds like the carbon monoxide and isocyanide above which were quite impossible on the old valence theories. The flexibility of the electron theory leads to peculiar valences such as univalent oxygen in the amine oxide and bivalent hydrogen.⁷ One of the interesting structures commonly accepted by the organic chemist is trivalent oxygen as found in the pyrilium ions.



This structure is not absolutely confirmed by the experimental evidence but it is a possible one. It would be a serious contradiction to London's theory as to the valence of oxygen. The organic chemists apparently have not realized that this oxygen is in a radically different state from oxygen in any other compounds except perhaps carbon monoxide.

The possibility of carbon forming a stable shell of six electrons has been mentioned. The structure



is a possible one for the compound $(\text{C}(\text{OR})_2)$ which Scheibler (20) claims to have prepared.

It will be seen from the foregoing that the electron theory of valence is capable of being brought into correspondence with an enormous number of experimental facts. The same may of course be said for the quantum theory of the atom. On the other hand it seems impossible at present to get rid of arbitrary postulates in either theory. These postulates must of course eventually become the same for both theories. It is evident at last that physicist and chemist are observing the same atom and the chemist may be confident that his observations will be given more consideration in theories of atomic structure in the future than they have been given in the past.

REFERENCES

- (1) BRIDGMAN: *The Logic of Modern Physics*, N. Y., Macmillan Co. (1927).
- (2) EDDINGTON: *Nature* **117**, 25 (1926)
- (3) BIRKHOFF. *Science* **65**, 147 (1927).
- (4) LEWIS: *Valence and the Structure of Atoms and Molecules*, N. Y., Chemical Catalog Company (1923)
- (5) HEISENBERG. *Z. Physik* **43**, 172 (1927).
- (6) TAYLOR. *Phys. Rev* **28**, 576 (1926)
PHIPPS AND TAYLOR *Phys. Rev* **29**, 309 (1927)
- (7) PARSON. *Smithsonian Inst. Pub* **65**, No. 11 (1915)
- (8) UHLENBECK AND GOUDSMIT *Nature* **117**, 264 (1926)
- (9) PAULI *Z. Physik* **31**, 765 (1925)
- (10) LEWIS. *Chem. Reviews* **1**, 231 (1925).
- (11) LONDON. *Z. Physik* **46**, 455 (1928)
- (12) HEISENBERG *Z. Physik* **41**, 257 (1927)
- (13) MAIN SMITH. *Chemistry and Atomic Structure*, London, Benn (1924).
- (14) GRIMM AND SOMMERFELD: *Z. Physik* **36**, 36 (1926).
- (15) WEISSENBERG. *Physik Zeit* **28**, 829 (1927).
- (16) LOWRY. *J. Chem. Soc* **123**, 822 (1923)
- (17) KONKYREV. *J. Russ. Chem. Soc* **52**, 17 (1920)
- (18) GADDUM. Private communication, from University of Missouri
- (19) HEIN. *Z. Elektrochem.* **28**, 469 (1922); *Z. anorg. allgem. Chem.*, **141**, 161 (1924)
- (20) SCHEIBLER *Ber* **60**, 554 (1927); **59**, 1022 (1927).

VALENCE FORCES AS CALCULATED FROM SPECIFIC HEATS AND ABSORPTION SPECTRA

DONALD H. ANDREWS

Department of Chemistry, Johns Hopkins University, Baltimore

Chemistry is concerned with the force of valence chiefly when the force is at the point of breaking down and permitting chemical action. It is therefore not to chemical but to physical properties that we must turn in order to get an idea of the nature of valence force when the chemical bond is in its really stable state. For example, in our concept of the thermal energy of a crystal as the vibration of the atoms restrained by elastic forces acting between them, there is illustrated one of the most important features of the stable chemical bond, namely its spring-like nature. Because of this relation between the chemical bond and thermal energy, crystals of organic molecules make an exceptionally good medium in which to study the bond in its normal state, for in these crystals practically the only forces which restrain the thermal motions of the atoms are the non-polar chemical bonds, those being many times stronger than the intermolecular forces. The exact nature of this relation can be most simply seen in the crystal lattice of diamond.

THE VALENCE FORCE IN DIAMOND

Diamond is believed to be made up of carbon atoms in a lattice such that each one is surrounded by four others so placed that lines drawn from the nucleus of an atom to those of its four neighbors pass through the corners of a tetrahedron. A carbon atom vibrating with thermal energy will oscillate about some mean position, restrained by these forces. Now it has been found by Einstein and Debye (1) that to account for the thermal energy of such a crystal lattice we must assume that each atom acts as a harmonic oscillator. This means that the force acting on it

must be such that when it is displaced from its mean position by a distance x , the force tending to restore it will be proportional to x , so that the value of the force can be expressed as

$$F = -kx \quad (1)$$

Thus if we measure distance in Ångstrom units (10^{-8} cm.), k will be the force in dynes acting on the atom when it is displaced one Å from its mean position of equilibrium. The distances through which the atoms move under the influence of thermal

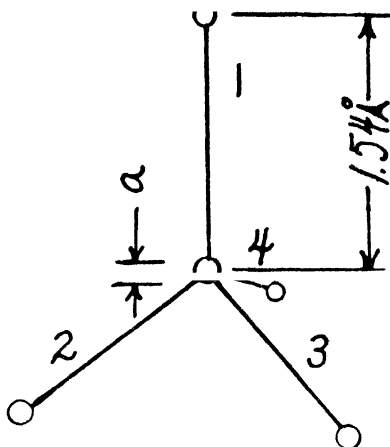


FIG. 1 FORCES ON CARBON ATOM IN DIAMOND

a = approximate amplitude of thermal motion at 700°C .

energy are about 0.2 Å so that this law of force should hold over distances of this order of magnitude.

We believe this restraining force to be actually the sum of the forces of valence acting on the atom. To produce a restraining force of a spring-like nature each valence force must therefore act like a spring. Moreover we find *diatomic* gas molecules also acting as harmonic oscillators, and from this it appears that a *single* valence force is like a spring such as pictured in figure 2. It is in equilibrium in a position such as shown in A. When the atoms are pushed together through a distance c then the force

repelling them is kc ; when they are pulled apart through a distance b , then the force attracting them together is $-kb$.

In diamond the force tending to restore a displaced atom to its equilibrium position is the vector sum of the four valence forces. Thus for a small displacement along the line of force 1, figure 1,

$$F = -(k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{3}k_4)x \quad (2)$$

because *force 1* will act in full, but the components of forces 2, 3, and 4 along the direction of *force 1* are only $\frac{1}{3}kx$ since the cosine of the angle between them, $70^\circ 32'$ is one-third.

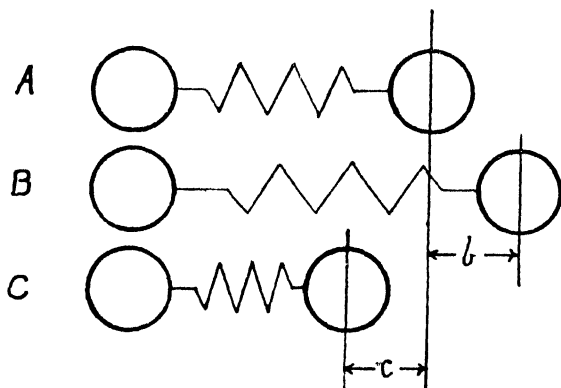


FIG. 2 SPRING-LIKE NATURE OF VALENCE FORCE

An atom oscillating under restoring forces of this kind will vibrate with a definite frequency. This frequency is related to the force constant k and the mass as follows:

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (3)$$

Thus if we know the mass of the atom m and the nature of the crystal structure in which it is located, together with the actual position of the valence forces acting upon it we can calculate the values of the forces from the frequency (ν).

The frequency is related to a number of other physical properties. In certain substances, light of the same frequency as the

vibration of the atom will be strongly reflected so that we can get the frequency from the so-called *rest strahlen* or residual rays of Rubens (2). Again in the infra-red absorption spectra there will generally be strong absorption of light of the frequency with which the atom is vibrating. These are *optical* methods of determining frequency.

The frequency is also related to the melting point in a manner first pointed out by Lindemann (3),

$$\nu = 3.08 \times 10^{12} \sqrt{\frac{T_m}{MV}} \quad (4)$$

where ν is the frequency, T_m the melting point, M the gram molecular weight, and V the molecular volume. This relation is based on the idea that at the melting point of a monatomic crystal the atom must possess the energy just sufficient to make its amplitude of vibration equal to the average distance between the atoms. An amplitude greater than this disrupts the crystal and causes melting.

Finally the frequency is related to the heat capacity. Einstein (1) showed that the heat capacity C_v of N harmonic oscillators (N = Avogadro's number) is expressed by

$$C_v = R \frac{\left(\frac{\Theta}{T}\right)^2 e^{\frac{\Theta}{T}}}{(1 - e^{\frac{\Theta}{T}})^2} = R F_E \left(\frac{\Theta}{T}\right) \quad (5)$$

where R is the gas constant, T absolute temperature and Θ what may be called the converted frequency, i.e. the real frequency multiplied by the constant¹ β , so that it has the dimensions of temperature. This expression which we may call the Einstein function, F_E , has been modified by Debye (4) to take into account the elastic waves in a crystal lattice. His function we will write as

$$C_v = R F_D \left(\frac{\Theta}{T}\right) \quad (6)$$

¹ β is Planck's constant, h , divided by the gas constant per atom, k and is taken equal to 4.778×10^{-4} in the following calculations.

Since both these functions depend only on $\left(\frac{\Theta}{T}\right)$ it is possible to make tables similar to logarithm tables in which values of C_v are given for corresponding values of Θ/T . Such tables (5) have been used in making the calculations which follow.

For diamond it is found that with $\Theta = 1860$, equation 6 reproduces very closely the experimental values for the heat capacity. Dividing this Θ by β gives a value for the frequency ν of 3.89×10^{13} vibrations per second. Putting this in Equation 3 we find that the force constant is 1.17×10^{-2} dynes per Å. Since the force due to a single valence is one-half the total force according to equation 2, we can say that a carbon-carbon bond stretched one-tenth of an Ångstrom unit exerts a force of 5.8×10^{-4} dynes, the actual mechanical force of the chemical bond. The law of force, from which we have derived this value, will probably be quite closely applicable up to a stretching of 0.2 Å, but beyond that other considerations enter in to change it, as will be shown later.

THE VALENCE FORCES IN BENZENE

The problem of calculating the values of the valence forces in an organic molecule is much more complicated than in diamond. It is a question of getting the space arrangement of the atoms in the molecule, and of the molecules in the crystal, then analyzing the motions of the atoms into characteristic types, and finally ascertaining the frequency of each type of motion. To do this precisely involves the solution of some very difficult dynamical problems and requires more intimate knowledge of the mechanics of the molecule than we possess at present. It is, however, possible to make an approximate solution, giving values for the frequencies which should be fairly accurate, and from these the values of the different valence forces can be calculated with the help of equation 3, as was done in the case of diamond. The way in which this is carried out may perhaps best be illustrated by giving in detail the calculation for benzene.

From the analyses of the crystal structure of aromatic organic compounds with the help of x-rays, we believe the carbon atoms

in the benzene molecule to be arranged in a hexagon as is postulated by the customary organic chemistry formula. These hexagons are probably stacked together like plates into the crystal lattice, each ring being surrounded by neighboring rings somewhat as shown in figure 3.

In a crystal lattice like this if we analyze the possible different motions into certain characteristic types we may say as a first approximation that the heat capacity of the crystal will be equal to the sum of the heat capacities of each of the different modes of motion. Moreover, the heat capacity of each mode of motion, or more technically speaking, each degree of freedom will at any

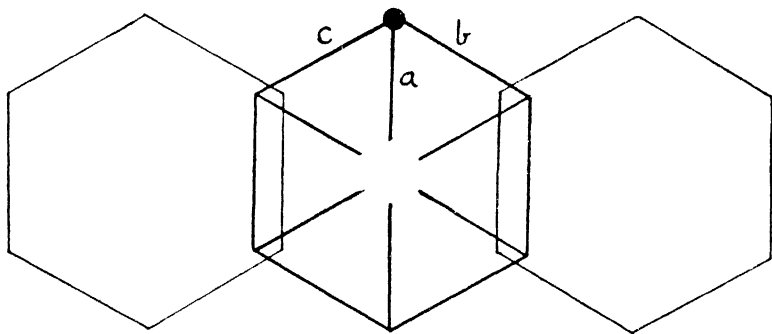


FIG. 3 VALENCE FORCES IN BENZENE

temperature depend only on the frequency of that motion and its value can be calculated from the frequency by means of equations (5) or (6). Thus if we find values for the frequency for the different types of motion and calculate the heat capacities, the sum of these should equal the experimentally observed heat capacity of the crystal. If this is the case, it is evidence that the analysis of the motion is correct and we can calculate the valence forces by means of equation 3 with a fair degree of confidence.

The first type of motion we consider is that of a single molecule in the lattice, moving as a unit, or as it will be termed hereafter, motion of the molecule as a whole. This type of motion occurs because the interatomic forces, i.e., valence forces, are, as will presently be shown, about sixty times as strong as the intermo-

lecular forces. It will have 6 degrees of freedom, three of translational and three of rotational vibration, and as a first approximation we can consider the frequency in them all to be the same.²

TABLE 1

The heat capacity of benzene, in calories per degree per mole

The observed values are from W. Nernst, *Ann Physik.* (4), **36**, 395 (1911).

TYPE OF MOTION $T^{\circ}K$	MOLECULAR $6 F_D \frac{(159)}{T}$	CU GROUPS			EXPANSION	TOTAL	
		1 $6 F_E \frac{(511)}{T}$	2 $6 F_E \frac{(1258)}{T}$	3 $6 F_E \frac{(2080)}{T}$		Calculated	Observed
10	0 2					0 2	0.3
20	1 7					1 7	1 7
30	4 0				0 0	4 0	4 5
40	6 1				0.1	6 2	6 6
50	7 5	0 0			0 2	7 7	8 0
60	8 7	0 2			0 4	9 3	9 4
80	9 9	0 8			0 7	11 4	11 1
100	10 5	1 9			0 9	13 2	12 9
120	11 0	3 1	0 0		1 1	15 2	14 4
140	11 2	4 3	0 1		1 4	17 0	16 0
160	11 4	5 4	0 3		1 6	18 7	17 4
180	11 5	6 3	0 5		1 8	20 1	19 0
200	11 6	7 1	0 9	0 0	2 1	21 7	20 5
220	11 6	7 7	1 2	0 1	2 3	22 9	22 7
240	11 7	8 2	1 7	0 2	2 5	24.3	24.8
260	11 7	8 7	2 2	0 3	2 7	25 6	27 0

To secure a value for these frequencies we can make use of equation 4 as explained above.³ Then using this value of the frequency in Debye's equation we can calculate the heat capacity

² The detailed reasoning for this analysis is given in a preliminary report on the specific heats of organic compounds by the author in the *Proceedings of the Royal Academy of Amsterdam* **29**, 744 (1926). See also *Communications from the Physical Laboratory, Leiden*, Suppl **56**, 16 (1926). It may be stated very briefly as follows. A single atom in a crystal lattice-like diamond can move in three directions that is, in the three dimensions of space, and is said to have three degrees of freedom because its position can be described by the three space coordinates. On the other hand, a molecule like benzene, which has a moment of inertia not possessed by an individual atom, can have energy of rotation. Its

due to the motion of the molecule as a whole. The values for this are given in the first column of table 1. Moreover by using these values of Θ we can with the help of equation 3 calculate the values of k , the intermolecular force. For a whole benzene ring this has the value of 5.5×10^{-5} dynes for a displacement of 0.1 Å. The force acting on a single CH group will be 0.92×10^{-5} dynes, so that it can be seen that it is very small compared with the C—C bond which is 58×10^{-5} dynes.

The next type of motion to be considered is that of the CH groups. In order to analyse this it is necessary to choose some arrangement of the valence forces in benzene and, as the double bond is still an uncertain quantity, it is assumed that each carbon atom has a bond similar to that in diamond, joining it to its neighboring carbon atoms and a third bond directed toward the center of the ring. The nature of this third centric bond is decidedly uncertain but there are some considerations which indicate that it is similar to the normal bond. For example, the heat of formation of benzene appears to be almost exactly the sum of six C—H heats of linkage and nine normal aliphatic C—C heats of linkage.⁴ With regard to the space arrangement of the

motion will have to be described by six coordinates like any rigid body, and it is therefore said to have six degrees of freedom.

Applying Equation 6 to an "atomic" crystal lattice we get

$$C_v = 3 RF_D \left(\frac{\Theta}{T} \right)$$

since each atom has three degrees of freedom and will be equivalent to three harmonic oscillators. For a "molecular" crystal lattice

$$C_v = 6 RF_D \left(\frac{\Theta}{T} \right)$$

since there will be both translational and rotational vibration. This implies torsional elastic waves, in addition to translational waves. This assumption is justifiable according to the best authorities on the subject.

³ The application of this equation to molecules will be described in a paper to be published soon in the *J Am Chem Soc*. It was presented at the St. Louis meeting of the American Chemical Society by J. C. Southard and D. H. Andrews.

⁴ Values taken from Eucken, Jette and LaMer *Fundamentals of Physical Chemistry*, N. Y., McGraw Hill (1925), p. 637 are as follows:

C—H	87 calories per mol.
C—C	70 calories per mol.
Calculated value for benzene will be $6(\text{C—H}) + 9(\text{C—C})$...	1152 calories
Observed heat of formation.....	1155 calories

CH groups, it is doubtful whether the atoms are in a plane or in a puckered hexagon, but since the difference between these two arrangements is less than the probable error of the calculation, we assume a plane hexagon. The motion characteristic of a CH group will probably be made up of *two* types of motion in the plane of the ring which will stretch directly some of the bonds, and, a *third* type of motion perpendicular to the plane of the ring which will not directly stretch any of them.

Considering the first type of motion for the CH group at the top of the ring in figure 3, motion along the line of the valence directed toward the center will have a force constant as shown in the following equation

$$k_a = k_1 + (k_2 + k_3) \cos 30^\circ$$

Taking all the valences the same as those in diamond, that is $k_1 = k_2 = k_3 = k_D$ then $k_a = 16.0 \times 10^{-4}$ dynes. Then, from equation 3, $\theta = 2080$.

In a system as small as the benzene ring we will not expect elastic waves to enter into the heat capacity to an appreciable extent so that the Einstein equation should be applicable. Employing θ with this equation, number 5, we find values for the heat capacity of this motion as shown in column 5, table 1.

In a similar way we find k_b for motion in the plane, perpendicular to the first type.

$$k_b = (k_2 + k_3) \cos 60^\circ = 5.85 \times 10^{-4}$$

From equation 3, $\theta = 1258$, and employing this in equation 4 the values of the heat capacity given in column 4 of table 1 are computed. There is no special reason for calculating the heat capacity for motion in these two particular directions, but they are representative of any direction on the plane and the choice of any others will not appreciably affect the result.

Motion perpendicular to the ring plane will, however, be opposed by much weaker forces. It will not stretch directly any of the bonds and, as has been seen, the intermolecular force opposing this type of motion is only $\frac{1}{6}$ th of the force of a C—C

bond. There is no evident means of calculating the frequency of this motion which, in a way, depends on the ease of bending of the bonds, a topic on which there has been much discussion, but little evidence. We have, however, one possible source of

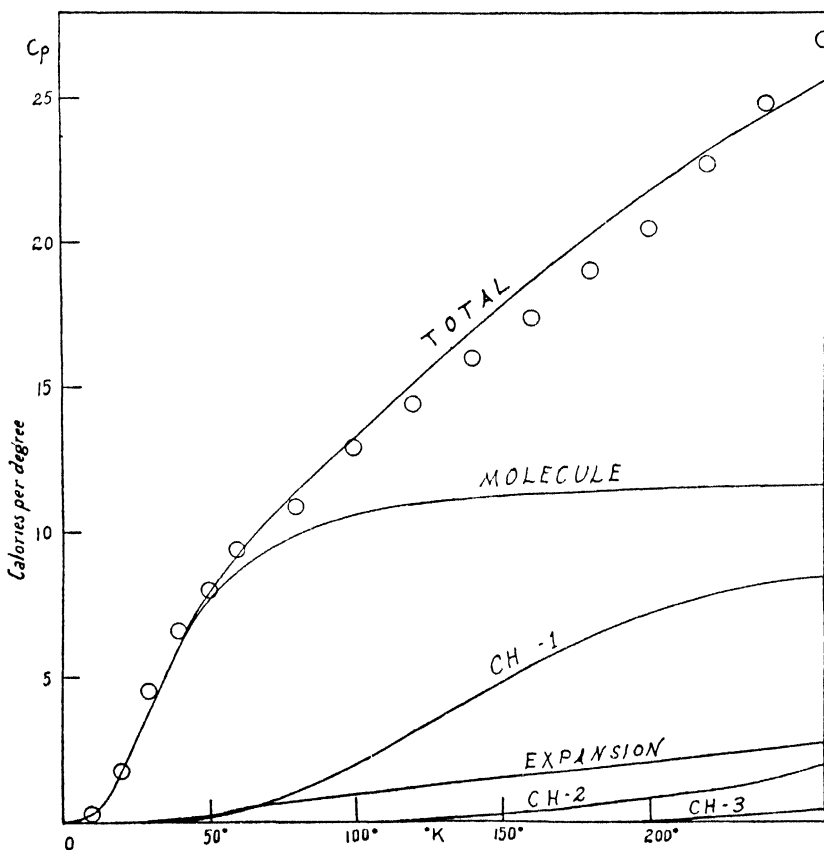


FIG. 4. THE HEAT CAPACITY OF BENZENE
Circles are values observed by Nernst

help. In the infra-red absorption spectra Ellis (6) has found a series of bands which seem to be quite definitely associated with the carbon-carbon bond. There are some grounds for believing that the frequency of the initial band of this series is that of

motion of CH groups perpendicular to the bonds rather than along them, from considerations of the electric moment of a carbon chain. If we employ the frequency of the bond $\theta = 511$ in the Einstein equation we get values for the heat capacity of this type of motion which are given in column 4, table 1. This will give the force for bending a bond as 0.32×10^{-4} dynes when the CH is moved 0.1 \AA . These three types of motion thus make up the three translational degrees of freedom for each CH group. The rotational degrees of freedom will probably be connected with motion of the H atom under the influence of the CH bond, the next topic to be considered.

The value of the force constant for the CH bond can be most easily secured from infra-red absorption spectra. There is a series of bands definitely associated with this bond with the initial band located at about 3μ . This gives a frequency of $\theta = 4280$ and a force of 4.78×10^{-4} dynes for a stretching of the bond of 0.1 \AA . Calculations from the Einstein Equation with this frequency show that motions of the hydrogen atom will not appreciably affect the heat capacity of crystalline benzene at any temperature.

The sum of the terms we have calculated should therefore give values of C_v for benzene. In order to compare these with the experimental values of C_p we have used an adaptation (7) of Nernst's equation for the heat absorbed due to the expansion of the crystal.

$$(C_p - C_v) = (C_v)_{\text{mole}}^2 \frac{T}{T_m} \times 0.0214$$

where $(C_v)_{\text{molecule}}$ is used because in view of the difference in the magnitude of the interatomic and intermolecular forces it appears that expansion of the molecule itself is negligible.

Adding the values calculated with the help of this equation to the sum of the other heat capacity terms we get values of C_p which we can compare with experiment. As shown in table 1 and in figure 4 the agreement is as good as could be expected in view of the approximate nature of the calculation. It appears that the choice of frequency for the molecular motion, and for

the motion of the CH group perpendicular to the plane of the ring, is at least qualitatively correct. About all it shows con-

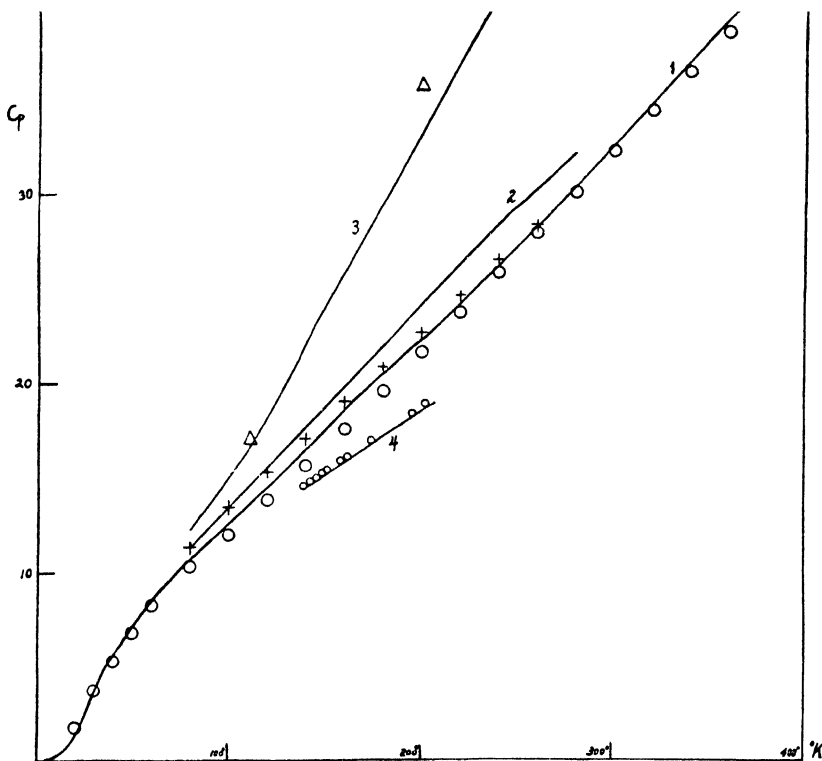


FIG 5 HEAT CAPACITIES FOR: \circ , 1, HYDROQUINONE; +, 2, GLYCEROL; Δ , 3, DEXTROSE; \circ , 4, METHYL ALCOHOL

The lines represent the calculated values. The symbols show the observed values taken from the data of (1) F. Lange, *Z. physik Chem* **110**, 351 (1924), Andrews, Lynn, and Johnston, *J Am Chem. Soc* **48**, 1274 (1926), D H Andrews, *J Am. Chem Soc* **45**, 1293 (1923); (2) Gibson and Giauque, *J Am. Chem Soc.* **45**, 93 (1923); (3) Simon, *Ann Physik* (4) **68**, 241 (1922); (4) Parks, *J. Am Chem. Soc* **47**, 338 (1925).

cerning the other frequencies is that they are so high that they do not affect the heat capacity appreciably. On the whole it seems to justify the choice of frequencies and the location of the

bonds, and shows that this is a fairly correct picture of the mechanics of the benzene molecule.

In a similar way the heat capacities of a number of other organic molecules have been calculated⁵ and the comparison with experimentally measured values is shown in figure 5. The same values for the force of the different normal bonds were used as in

TABLE 2
The force of valence

BOND	DYNES FOR STRETCHING 0.1 Å.	OBSERVED HEAT OF LINKAGE	ϵ_D
		kgm. cal./mole	Å units
C—C			
Stretching	5.85×10^{-4}	81.6	0.44
Bending	0.32		
Intermolecular force per CH group in benzene	0.092		
H—H	4.60	101.0	0.55
H—N	5.62	98.0	0.48
H—C	4.78	90.0	0.51
C—Cl	1.88	81.0	0.77
C—Br	1.89	39.7	0.54
C—I	2.08	40.0	0.52
Cl—Cl	3.30	58.5	0.49
Cl—I	2.36	49.6	0.54
Br—Br	2.48	45.2	0.50
I—I	1.68	35.2	0.54

Note: For the sources of most of the above values see Bates and Andrews, loc. cit. Certain values have been revised somewhat since first published.

the case of benzene. It has not been possible up to the present time to extend the study far enough to include the more abnormal bonds such as the tertiary C—H, and C=C, but it is hoped that studies of heat capacity and absorption spectra which are now being carried on will make these methods of use in solving the problems of chemical activity.

⁵ A slightly different method of calculation has been proposed by E. O. Salant, Proc. Nat. Acad. Sci. 12, 334, 370 (1926). For further details of the above method of calculation see the paper by the writer in the Proceedings of the Royal Academy of Amsterdam, loc. cit.

THE CHANGE OF VALENCE FORCE WITH STRETCHING OF THE BOND

It is interesting to compare with each other the various values of the force of valence which can be secured. The values calculated above, together with a number calculated by Bates and Andrews (8) are shown in table 2. As these authors have pointed out, it is remarkable how close together the values lie for the different non-polar bonds. It seems to indicate a common mechanism which is very nearly the same in all the bonds.

There is other striking evidence of similarity in the behavior of the bond when stretched to distances greater than atomic di-

27

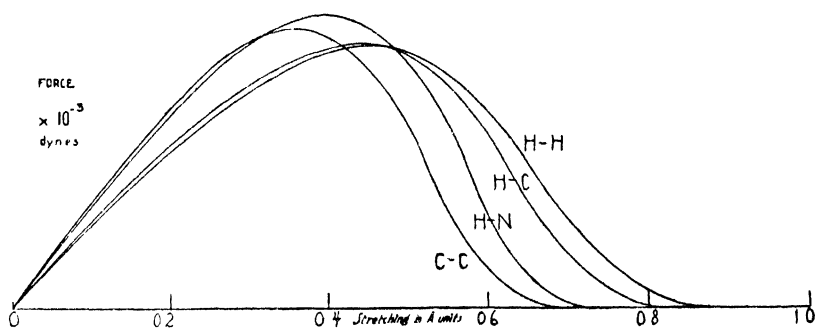


FIG. 6. VALUES OF THE VALENCE FORCE WHEN THE BOND IS STRETCHED

ameters. We know from the heat of linkage the value of the work done in pulling two atoms apart and severing the bond completely. Now if the law of force $F = -kx$, which holds for a stretching of 0.2 Å, should continue to hold for greater distances, we could calculate the amount of stretching which would just perform the amount of work equal to the heat of linkage. Values for this distance, x_D , as calculated by Bates and Andrews are given as x_D in table 2, column 3. If this were the case we would expect that the force would suddenly cease at this point, and the atoms would fly apart without further restraint. In all probability this is not the case. We have evidence from the term

values of absorption spectra that even at distances of 0.3 \AA in $\text{H}-\text{H}$ and $\text{C}-\text{H}$ the force is considerably less than that required by Equation I, and it probably begins to decrease as the bond is stretched over 0.4 \AA . But in the fact that on the basis of Table 2 in all cases about the same amount of stretching is required to perform the work equal to the heat of linkage, we have evidence of the great similarity of all non-polar bonds.

This same relation is shown graphically in figure 6. Following the general shape of the force curve for $\text{H}-\text{H}$ as calculated from spectral term values,⁶ the force curves for the different bonds have been drawn so that the area under each one is equal to the heat of linkage. The portions of the curves beyond 0.4 \AA . have little more than qualitative significance because we do not know whether atoms pulled apart may not even be repelled after a certain distance is reached. This does show, however, that it is quite probable that a stretching of something a little less than 1 \AA is involved in most chemical processes. If a chemical reaction involved performing work against the force of the bond for distances greater than this, the heat of activation would be so large that action under most conditions would be impossible.

A study of the chemical bond from facts of this sort shows what the mechanical nature of the bond probably is, though it does not explain the mechanism, electrical or magnetic, producing the mechanical force. We hope, nevertheless, that it may serve as a basis for inquiring further into the problems more closely related to chemical activity and to the thermal decomposition of molecules.

REFERENCES

- (1) EINSTEIN *Ann Physik* (4) **22**, 180 (1907)
DEBYE *Ann Physik*. (4) **39**, 789 (1912).
- (2) RUBENS AND NICHOLS *Ann. Physik* **60**, 418 (1897).
- (3) LINDEMANN: *Physik Zeit* **11**, 609 (1910). For the value of the empirical constant see W. Nernst, *Vorträge über die kinetische Theorie der Materie und der Elektrizität*, Göttingen, (1914)

⁶ This was calculated with the help of the equation for an anharmonic oscillator given in Max Born: *Vorlesungen über Atommechanik*, Berlin, Springer (1925), p. 78. A treatment of this problem with the help of the new quantum mechanics has just been published by Sugiura *Z Physik* **45**, 491 (1927).

- (4) Loc cit. See also Born and Kármán: *Physik. Zeit.* **14**, 15 (1913).
- (5) NERNST: *The New Heat Theorem*, N. Y., Dutton (1927).
MACDOUGALL: *Thermodynamics and Chemistry*, 2nd ed, N. Y., Wiley (1926).
- (6) ELLIS: *Phys Rev* **23**, 48 (1924); **27**, 298 (1926)
MARTON: *Z physik Chem* **117**, 97 (1925) and *Bell: J Am Chem Soc.* **47**,
2811 (1925) for studies on infra-red absorption spectra of organic
compounds.
- (7) NERNST: *Ann. Physik.* (4) **36**, 395 (1911)
- (8) BATES AND ANDREWS: *Proc Nat. Acad. Sci* **14**, 124 (1928).

THE RELATION OF SHARED ELECTRONS TO POTENTIAL AND ABSOLUTE POLAR VALENCES

WILLIAM ALBERT NOYES

Chemical Laboratory, University of Illinois, Urbana

Dalton proposed his atomic theory of the constitution of matter and of the combinations between elements about 125 years ago. Since then the advances in chemistry and physics have been, almost continuously, in agreement with the hypothesis that matter is discontinuous in what may be called its fine structure. For some time, now, there has been a growing belief that energy is also discontinuous and that matter and energy are interconvertible. When we find a theory of this sort, which has been continuously and increasingly useful for more than a century, it is wise for us to consider whether new hypotheses which may seem, at first, incompatible with the older ideas, may not be reconciled with them.

With this principle in mind, the endeavor will be made to show that the electron theory of valence proposed by G. N. Lewis (1) may be reconciled with the original theory of valence proposed 75 years ago, and with the positive-negative theory proposed by Arrhenius 40 years ago (2) and applied to compounds not usually thought of as forming ions, a few years later (3). This will be done with the use of three rather simple principles:

1. In the majority of chemical reactions, a pair of electrons forming a covalence remains with one of the two atoms which separate (4).

2. In every ion, whether simple or complex, there is some atom for which the sum of the covalences plus the number of unshared electrons is greater or less than the number of valence electrons for that atom (5).

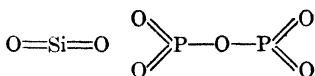
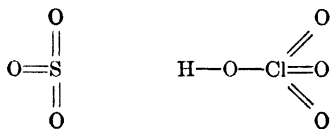
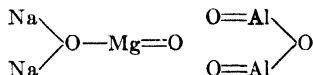
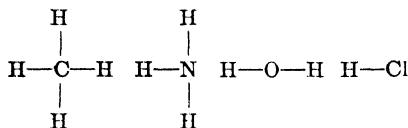
3. In every compound having only covalences, the number

of covalences plus the number of unshared electrons is equal to the number of valence electrons for each atom (5).

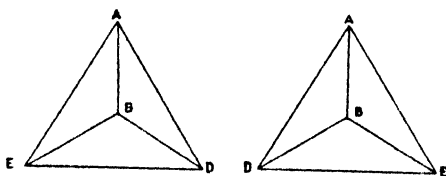
The second and third principles are partly implied in many things which have been said by G. N. Lewis and Langmuir but they have not stated it clearly in the above form and they and others have not always recognized that valence is a property which changes by *whole* numbers. These numbers are not dependent on the stability of the compounds or on electrical moments of compounds which may render some covalences potentially polar (see below). They have not made the sharp distinction between polar and non-polar valences on the basis that at a given instant any valence must be either an ionic valence or a covalence. They have used the term "semipolar valence" not as something which is half ionic and half covalence but as a single covalence in which there is an unusual electrical moment. The practical difference, here, is not very great and is partly a question of nomenclature rather than a difference of opinion about what may be actually present in the molecule.

These principles, if accepted, not only furnish a reconciliation between the older and newer theories of valence but they also account for the variable valences which gave chemists so much trouble 50 or 60 years ago.

The original idea of valence, as developed by Frankland, Couper and Kekulé, is that each atom has one or more unit powers of attraction for other atoms. This idea is still true, as will be seen in what follows. I think of it as one phase of the discontinuity which lies at the foundation of all natural phenomena. It is expressed by the following well known formulas:



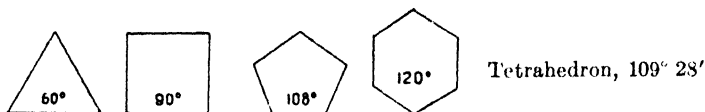
Pasteur, LeBel and van't Hoff added the idea that each atom is held to a central atom having four valences, in such a manner that right handed and left handed arrangements are possible, with sufficient stability to prevent the easy interchange of the positions of the groups:



Van't Hoff furnished the further idea that doubly bound carbon atoms do not rotate freely about the axis of the double union, while singly bound atoms may do this:



Baeyer suggested that rings containing 5 or 6 atoms are more stable than those having only 3 or 4 atoms:



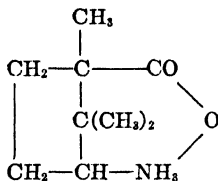
Arrhenius gave us the modern idea of polar valences when he assumed that the ions of an electrolyte are more or less completely separated in solutions and exist as independent molecules, having positive and negative charges. Some of the followers of G. N. Lewis have carried this thought so far as to deny that there is a genuine valence between such ions. They seem to have overlooked three important facts:

1. Every negative ion must have a positive ion in the immediate neighborhood. The attractions between the ions hold them together in this way.

2. If hydrochloric acid or sodium chloride is vaporized, the atoms are held together in pairs in the gaseous state. Whether

the atoms are held together by a polar valence or a covalence is not certainly known.

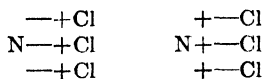
3. Aminolauronic acid is held in a stable configuration,



in solution, by the polar valence between the oxygen atom and the ammonium group (6).

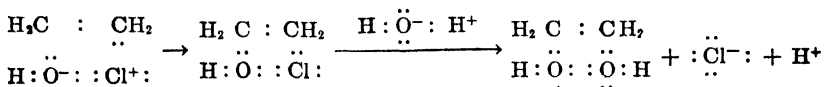
About 25 years ago, J. J. Thomson proposed the first definite electronic hypothesis (7) to account for chemical combination. He assumed that after the transfer of an electron from one atom to another the atoms were held together by the static attractions between them. This idea has been made the basis for a discussion of chemical reactions by Falk and Nelson and a number of others, including myself. I was finally convinced that the pair of electrons forming a covalence are held by the two atoms in common, as supposed by G. N. Lewis, by the following considerations:

1. The theory that an electron is transferred from one atom to the other implies the possibility of the electromers:



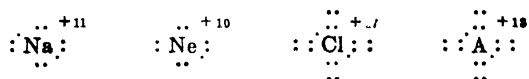
A fruitless search, continued for nearly ten years, with the aid of several able assistants, failed to disclose the existence of such electromers.

2. Hypochlorous acid adds to ethylene forming ethylene chlorohydrin and the latter hydrolyses to glycol and hydrochloric acid. It is evident that in the first reaction the chlorine atom adds to the carbon in the positive form with only six electrons (8) and leaves it in the negative form with eight:

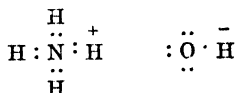


This is easily understood if the pair of electrons is shared by the chlorine and carbon in the chlorohydrin but it is not readily explained if the electron is transferred completely from one atom to the other.

Atomic ions of the first two periods of the periodic system have the same electronic shell as that of the nearest noble gas. They differ from the noble gas in the positive charges of their nuclei. The positive charge of the nucleus of the sodium ion is one unit greater than that of the nucleus of neon. The positive charge of the nucleus of the chloride ion is one unit less than that of the nucleus of argon:

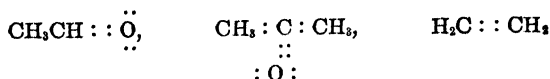


In complex ions there is some atom which has more or fewer electrons than those required for atoms having only covalences. Nitrogen has 5 valence electrons; ammonium has only four pairs of shared electrons and is a positive ion. Oxygen has 6 valence electrons; the hydroxide ion has 6 unshared electrons and one pair of shared electrons. It is a negative ion.

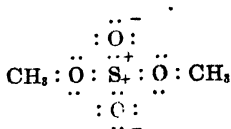


The noble gas structure of ions seems to be closely related to their semi-independent existence in solutions (9).

There seem to be two different kinds of double unions; double covalences, as in aldehydes, ketones, ethylene, etc.,



and a combination of a covalence and a polar union, as in dimethyl sulfate,



Here the sulfur atom, which has six valence electrons, has only four covalences and two positive polar valences. The oxygen atoms, above and below, have one negative polar valence and six unshared electrons. The oxygen atoms connected with the methyl groups have two covalences and four unshared electrons, in accordance with the rule. It is to be noted that the semipolar oxygen atoms show no more tendency to form covalences by sharing their electrons than chloride ions do. They seem to be genuine half ions, and resemble the noble gases in the same manner as other ions.

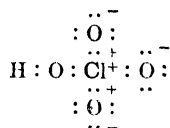
The point of view here presented furnishes a satisfactory reconciliation between the formula proposed by G. N. Lewis for the sulfate ion and the older view that oxygen is bivalent in nearly all of its compounds and that sulfur has a valence of six in sulfur trioxide and sulfuric acid. It receives strong support from the "parachor" relations established by Sugden (10) and from the optical activity of sulfinic compounds (11).

T. W. Richards suggested that some of the phenomena of chemical combination may be accounted for by the hypothesis of compressible atoms. The diminution in volume occasioned by the transfer of the valence electron of a metal to a halogen or other atom, completing its octet and the more compact volumes of molecules having octets completed by sharing electrons or drawn together by semipolar unions furnish a picture somewhat resembling that of Richards but much more definite than this.

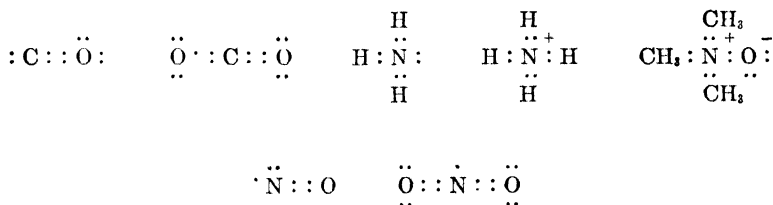
The distinction between the polar valences of ions and the covalences of other unions is so sharp that the author has proposed (12) to restrict the term "polar valence" to the designation of substances which are ions or semi-ions. In advocating such a nomenclature it is clearly recognized that some molecules have an electrical moment at the covalence and may separate, momentarily, into ions when reacting with other molecules. It will contribute greatly to clarity and definiteness in our discussions if such covalences are called *potentially* polar instead of merely polar.

The theory of G. N. Lewis, combined with the rule for shared and unshared electrons, gives a very satisfactory explanation of the

variable valences which caused so much trouble to chemists 50 or 60 years ago. Thus, a chlorine atom is univalent toward hydrogen, with either a polar valence in the chloride ion or, possibly, a covalence in the molecule, $\text{H}:\ddot{\text{Cl}}:$. In sodium chloride the polar valence seems to be retained in the crystal. In perchloric acid the chlorine atom has a valence of seven, consisting of 4 covalences and 3 polar valences:



The valences of carbon in carbon monoxide and carbon dioxide and of nitrogen in ammonia, ammonium, the amine oxides and the various oxides of nitrogen are easily explained in the same manner:



These principles give a good explanation for the reaction between nitric oxide and nitrogen trichloride, which gives nitrosyl chloride, nitrous oxide and chlorine, with nitrogen dichloride as a probable intermediate compound. This reaction has been studied for two years past and is the subject of a separate communication (13).

It should be noted that the discussions and illustrations of this paper have been confined to hydrogen and the first two periods of the chemical elements. It seems probable that other considerations are involved as the shell of electrons increases in size. G. N. Lewis has considered some elements having higher atomic weights.

REFERENCES

- (1) LEWIS: J. Am. Chem. Soc. **38**, 762 (1916).
- (2) ARRHENIUS: Z. physik. Chem. **1**, 637 (1887).
- (3) HOFF, VAN'T: Z. physik. Chem. **16**, 411 (1895).
NOYES AND LYON: J. Am. Chem. Soc. **23**, 463 (1901).
STIEGLITZ: Ibid. **23**, 796 (1901).
LAPWORTH: J. Chem. Soc. **79**, 267 (1901).
- (4) NOYES: J. Am. Chem. Soc. **45**, 2959 (1923).
- (5) NOYES: Z. physik. Chem. **130**, 323 (1927); Bicentenary No., Proc. Am. Phil. Soc. **66**, 302 (1927).
- (6) NOYES AND POTTER: J. Am. Chem. Soc. **37**, 189 (1915).
- (7) THOMSON: Phil. Mag. **7**, 237 (1904).
- (8) NOYES AND WILSON: J. Am. Chem. Soc. **44**, 1630 (1922).
BALARD: Ann. chim. phys. (2) **57**, 272 (1834).
- (9) NOYES: Proc. Nat. Acad. Sci. **13**, 379 (1927).
- (10) SUGDEN: J. Chem. Soc. **125**, 1177 (1924); **127**, 1528 (1924); **131**, 1173 (1927).
- (11) HARRISON, KENYON AND PHILLIPS: J. Chem. Soc. **129**, 2079 (1926).
- (12) NOYES: Z. physik. Chem. **130**, 329 (1927).
- (13) NOYES: J. Am. Chem. Soc., **50**, 2902 (1928).

A PRAGMATIC SYSTEM OF NOTATION FOR ELECTRONIC VALENCE CONCEPTIONS IN CHEMICAL FORMULAS

HARRY SHIPLEY FRY

Department of Chemistry, University of Cincinnati, Ohio

This paper presents for consideration some questions and points of view bearing upon the methods of indicating within chemical formulas electronic conceptions of valence now in vogue.

Alfred W. Stewart (1) in his latest volume, "Recent Advances in Organic Chemistry," states that "The greatest problem before Organic Chemists at the present day is the application of modern electronic views to the salient phenomena among the reactions of carbon compounds." If we accept this dictum, we are immediately confronted with two pertinent questions. First, in what manner and to what extent should the chemist's graphic or structural formulas be modified so as to conform rationally with current notions of electronic valency? Second, will such modifications elucidate or will they confuse the fundamental purpose of a structural formula, which, as all chemists agree, should not only present a picture of the number, the kind, and the arrangement of the atoms within the chemical molecule but also clearly indicate and correlate the manifold chemical reactions displayed by the molecule.

The import of both of these perplexing questions has been fully sensed and expressed by John A. Cranston (2) in the opening paragraph of an address entitled "Symbols for Electronic Bonds" in these words:

The development of the electronic theory of valency has not been accompanied by an expansion in symbolism adequate to express without ambiguity the large number of ideas that exist concerning the bonds that unite atoms together in a molecule. The horizontal line which originally indicated a chemical bond, the dot, and the plus and minus

Langmuir conception of electronic shells than to the Bohr conception of electronic orbits. A possible reason for this attitude may be the fact that the cubical octet conception is primarily and fundamentally an outgrowth and a pictorial elaboration, in terms of atomic structure, of the original conception of Abegg and Bodlander (9), who, in 1899, stated that atoms display different kinds of valency termed "normal" and "counter" valency of opposite polarity according as they are united with electropositive and electronegative atoms. Furthermore, it is a question in the minds of many chemists whether or not the Lewis-Langmuir valence conceptions and notations possess any more, or perhaps as much, truly chemical significance as that which characterizes the Abegg and Bodlander system.

Now when we come to consider the actual chemical properties of atoms and molecules, no matter what attempts may be made to explain valency by an electronic pictorial notation, all that the chemist knows pragmatically about the valence of an atom may be embodied in the simple fact that if (n) be the empirically determined valence of a given atom, that atom may function in ($n + 1$) different ways. This has been fully illustrated in the numerous publications (10) of the interpretations of a great variety of chemical reactions. This is also a direct and simplified modification of the conceptions of Abegg and Bodlander and it does not require the amplification and entailed ambiguity necessarily encountered when the cubical octet or other systems of electronic valence notation are imposed upon structural formulas. It enables us to correlate in a simple fashion the several ways in which atoms and radicals react, positively, negatively, and amphotERICALLY, in strict conformity with the actual chemical behavior of the molecule as amply illustrated by the well established chemical reactions actually being dealt with such as ionization and electrolysis, hydrolysis, and oxidation-reduction processes. In fact, practically all chemical phenomena may be classified under these types.

This idea and procedure may be termed "chemical pragmatism." It takes issue with attempts to embody within the structural formulas of the chemist's molecule the diverse meta-

physical hypotheses which are concerned chiefly with questions relative to the constitution of the atom and the disposition of its valence electrons. Of course, the more chemists know about the constitution of the atom, the more fully can they explain the chemical properties of the atom *per se*, but we must not lose sight of a paramount philosophical truth, developed by Immanuel Kant in his Critique of Pure Reason, namely the antinomies of pure reason which warn us that whenever assumptions or hypotheses are based upon premises that lie beyond the territory of sensation experience, that is, can not be observed or demonstrated through tangible experimentation, absolutely contradictory conclusions are bound to follow.

That this state of affairs is all too prevalent is attested by the previously noted plea of Cranston (2) for the adoption of a rational system of electronic valence notation. Not only the readers of journals, but some editors sense it. For example, an editorial note published in the Journal of the Society of Chemical Industry states that

Discussion of alternate polarities and kindred topics is of great importance, and we hope the letters we have published in these columns have cleared away a considerable mass of misapprehensions. We are, however, obliged to wait for a period before dealing with the subject again; *a proportion of our readers fails to understand the whole of the arguments without a mental effort which is made unwillingly.* (Author's italics.)

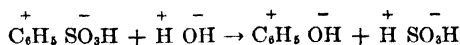
It is also evident that the authors of these arguments also frequently fail to comprehend one another. If you doubt this statement, just call to mind the arguments in an electronic valence discussion at any chemical society meeting. Two advocates of one and the same system of electronic valence notation will be diametrically opposed in their interpretation of the electronic character of a given atom or radical of a molecule: one says it is positive; the other claims it to be negative; and the chances are very strong that a third advocate of the same theory will arise and maintain that it is neither positive nor negative but non-polar because the electron pair is just midway in its location between

two atoms. Arguments on atomic structure then follow to show why the electron pair is more the property of one atom than of another and, in the end, all are wondering, both polar and non-polar advocates, "how many angels can dance on the point of a needle." Such disputes can be settled only by direct reference to the actual chemical behavior of the substance whose electronic formula is under discussion, and *only that formula should be assigned to the molecule which is in conformity with its particular behavior at that particular time.* As is well known, a given substance may be an electrolyte or polar compound in one solvent or it may be a non-electrolyte or a non-polar compound in another solvent. Temperature, concentration, and catalytic agents are other conditions which completely alter the character and reactivity of the molecule. *It is therefore impossible to state that a given substance is specifically a polar or a non-polar compound.*

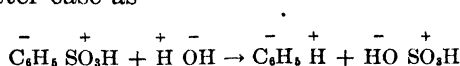
The electronic valence skeleton in the chemist's closet embodies many bones of contention, but its spinal column is the arbitrary distinction that has been set up between polar and non-polar compounds. W. A. Noyes (11) once stated that "The difference between polar and non-polar valencies is one of degree and not of kind." G. N. Lewis (12) has stated that "even a symmetrical molecule like that of H_2 or I_2 may from time to time become polarized in one direction or the other, as a consequence of the disturbance due to thermal motion. . . . In other molecules some displacement of electrons may occur without full ionization." Julius Stieglitz states that "Polarity exists wherever there is a difference in charge, positive and negative, between two atoms, and it is so valuable a conception in the treatment of organic compounds, non-electrolytes, that it must be clearly insisted upon. Again, just recently, Tanasescu (13) has emphasized the fact that chemists, notably those of the American school, influenced by data relative to the electronic conception of matter, that is of the constitution of the atom *per se*, have not considered the totality of the reactions of the organic molecule but rather have preferred to show with the greatest precision possible the nature of the affinities between atoms. Tanasescu deems it preferable to regard electronic conceptions of valence as independent of con-

ceptions relating to the inner nature of valency and maintains that, in the search for the specific mechanisms of atomic unions as complex as those of the organic molecule, *only a theory of polarity much more simplified will be of real interest and success.* All of this is dependent upon correlations which should be based primarily and fundamentally upon the actual chemical reactions of the organic molecule. Tanasescu further states that "even if the organic molecule is not capable of furnishing ions, as commonly conceived, we can always consider it as being susceptible to dissociation into atoms or radicals charged with quantities of electricity of equal and opposite sign."

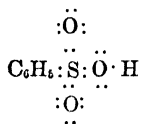
These points of view were originally advanced by Fry (4) in his various papers and monograph on the electronic conception of positive and negative valence and the constitution of benzene, and in his conceptions of *electromers* and *electronic tautomerism*, which afforded an explanation of the Crum Brown-Gibson benzene substitution rule and the *simultaneous* formation of ortho, meta, and para substituted derivatives of benzene (14). A given derivative of benzene, for instance phenylsulfonic acid (15) $\text{C}_6\text{H}_5\text{SO}_3\text{H}$, may react either as the electromer $\text{C}_6\text{H}_5^+\text{SO}_3^-\text{H}$ or the electromer $\text{C}_6\text{H}_5^-\text{SO}_3^+\text{H}$. The particular reaction occurring depends wholly upon the specific experimental conditions under which the reaction takes place. Thus, for instance, hydrolysis of phenylsulfonic acid *in alkaline solution* yields phenol and sulfurous acid while in *acid solution*, the products are benzene and sulfuric acid. Entirely apart from the question of whether phenylsulfonic acid is a polar or a non-polar compound, it reacts in the former case in strict conformity with the simple pragmatic polar valence formulas notation



and in the latter case as



Furthermore, there is nothing whatever in the shared electron octet polar valence system of notation for phenylsulfonic acid which Lewis (16) represents by the cumbersome formula



to indicate or to correlate directly its actual chemical behavior with its electronic structure, much of which is quite extraneous to the specific reactions in question. Lewis virtually begs this question in the following words:

It is no longer necessary to consider this substance as a mixture of two tautomeric forms [electromers], in one of which the phenyl group has a unit positive charge, and in the other of which it has a unit negative charge. . . . We see that the bonding pair between phenyl and sulfur may be shifted toward the one or the other, and when the molecule is broken at this point, the phenyl group, if it retains possession of the bonding pair, will combine with hydrogen ion but if it loses possession of the bonding pair will combine with hydroxyl ion.

Furthermore, this statement of Lewis may be regarded as a direct admission that phenylsulfonic acid does react under the different conditions noted as *two distinct tautomeric forms*, for if, in one case, the electron pair is in one position in the molecule and if, in the other case, the electron pair occupies a different position in the molecule, he is then undeniably dealing with two distinct electronic formulas, electromers, which are most simply and pragmatically represented by the polar valence formulas $\text{C}_6\text{H}_5\text{SO}_3^+\text{H}^-$ and $\text{C}_6\text{H}_5\text{SO}_3^-\text{H}^+$. In substance, Lewis attempted to paint a shared electron picture in one structure of two distinct tautomeric forms which Fry (17) earlier defined as the electromers or the electronic tautomers of phenylsulfonic acid. This is cited as only one of many illustrations wherein the simple pragmatically employed polar valence formula notation has been sacri-

ficed for the more complicated and abstruse shared electron and octet notation.¹

Some critical students have asked if this is not partly due to a current fashion or the common custom of courting popularity by adopting the latest ideas. I am reminded, in this connection, of a bit of advice received from a well known former advocate of the simpler electronic polar valence notation who urged me to drop the plus and minus notations and, as he expressed it, "get on the Lewis-Langmuir band wagon." Pragmatic chemists are not naturally disposed to discard the violin for the saxophone, even though both instruments may be profaned. Moreover, pragmatic chemists are not going to subscribe to a theory because it is fashionable or because of the psychological truth that "if you repeat something often enough, loudly enough, and blandly enough, eighty per cent of the world will believe it," regardless of its intrinsic merits or demerits.

All of this brings us back again to that dry bone of contention, namely the distinction between polar and non-polar compounds. The necessity for such a distinction does not concern pragmatic chemists, several of whom have been quoted in this paper. Further, in this same connection, James Kendall (18) has recently written two very significant statements in his admirable textbook on Inorganic Chemistry. First, "it is immaterial to us at this stage whether the Lewis-Langmuir conception of electron shells or the Bohr conception of electron orbits is finally accepted." Second, "Strictly speaking, the distinction between polar and non-polar substances is not fundamental but one of degree only." In the same vein, Professor K. George Falk, (19) one of the original advocates of the simpler electronic polarity formulas, writes "While it must be admitted that the octet theory is interesting to study, fascinating to work with, and permits of ready popular

¹ For more recent direct experimental evidence for the existence of electromers, "Two Forms of Ortho-Nitro-Toluene," see Clark and Crozier, Transactions of the Royal Society of Canada, 3rd series, Vol. XIX, 157 (1925). See also the survey by Clark and Carter, *ibid.*, Vol. XXI, 323 (1927), on "The Replaceability of Nitro Groups from the Nucleus of Various Aromatic Compounds," one hundred and ten in number.

presentation, at the same time the question may well be asked whether this theory as such adds anything real to an understanding of chemical reactions." While "the use of shared electrons for explaining chemical reactions possesses apparent advantages in some cases it would seem as if more definite postulates as to the positions of the shared electrons should be available before the formulation will possess real significance. *It may therefore be questioned whether the octet theory or the shared electron theory at the present time contributes any essential relation to the simple electron transfer valence view;*" that is, to the plain polarity valence formula notation. (Author's italics.)

Julius Stieglitz also strongly advocates the employment of a pragmatic system of electronic polar valence notation in the formulas and reactions of compounds. Quotations from a recently received letter are quite pertinent to the issue. Professor Stieglitz states:

Like yourself, I would insist on emphasizing polarity in organic (and inorganic) compounds whenever the actual chemical behavior of compounds indicates polarity. I believe, however, that a great deal of confusion would be avoided if we could agree to distinguish the polarity resulting from completely transferred electrons as *ionic polarity* and the polarity resulting from the definite approach of electron doublets to given atoms as *doublet polarity* or simply *polarity*. While in many cases the two may approach each other (as in acetic acid but not in sodium acetate), nevertheless their average behavior makes so profound an experimental difference that it would be a mistake, I believe, not to recognize the difference. Organic chemistry itself would be benefited thereby for instance in the matter of *tautomeric ions* of certain salts. Besides, failing to recognize and express the difference, we let outsiders condemn the whole question of polarity in organic compounds, because they think we mean ionic polarity in all cases when we certainly have something quite different in mind for alkyl halides, carbonyl derivatives, benzene derivatives, etc. Further, I think it is important to differentiate the two types in our symbols. I have urged a convention that the charges on ions be put in heavy type so that we could use ordinary type to indicate polarity in organic compounds, which is so essential for the modern theories of organic chemistry and for teaching much of it from the point of view of scientific theory

rather than of empirical rules. . . . It would do away at once with a great deal of misunderstanding on the part of physical chemists, which is the chief obstacle in the way of progress in this subject in our own country. It was on account of the lack of such a convention that, for instance, Lewis misread my "The Electron Theory of Valence as applied to Organic Chemistry" published in 1922. It stated explicitly at the outset that the Lewis-Bohr doublet theory of union of atoms is accepted as far as the nature of the bond is concerned . . . but insisted on polarity resulting from the definite approach of doublets to certain atoms. Because the ordinary plus and minus signs were used, the article was read by physical chemists, including Lewis, as if ionic polarity were represented.

Stieglitz continues further

As to the validity of our views on definite polarity in organic compounds as well as in many inorganic compounds supposed to be non-polar, I believe there can be no doubt in the minds of men who have actually worked with such substances and are familiar with their behavior. In the inorganic field, the striking difference in behavior of phosphorus trichloride and nitrogen trichloride are consistently and simply explained from the same point of view, Lewis notwithstanding. It is needless to say, in conclusion, that the theories of atomic structure would lead us to expect polarity as the result of the differences in positive kernels of atoms, which will lead to different forces of attraction for doublets. J. J. Thomson (20) develops this relation for methyl chloride in his "The Electron in Chemistry." In benzene derivatives, C_6H_5X , the influence of X, as X^+ or X^- is the key to the understanding of their behavior, a fact which, in its naked simplicity, you were the first to recognize and emphasize. I believe they act simply by attracting or repelling the bonds of the ring carbon atoms and have accumulated further evidence to that effect.

Let me conclude by an appeal to your sense of admiration for simplicity by reemphasizing the opinions of the authors I have been privileged to quote, namely that *pragmatic chemists* are more vitally concerned with the adoption of a system of electronic valence notation which is primarily designed to indicate and correlate the actual chemical behavior of the atoms and the radicals of the molecule as displayed under specific experimental conditions

than with the imposition upon structural formulas of an electronic system of notation which is complicated by metaphysical speculations involving the unsolved problems of the constitution of the atom and the disposition of the hypothetical electron shells or the elusive electron orbits of its valence electrons.

Please do not regard these pragmatic points of view relative to questions of atomic structure and to the elimination of octet and electron doublet valence notations from the structural formulas of organic and inorganic compounds as derogatory or prejudicial. Their purpose is to stress a growing opinion among pragmatic chemists that much of the modern speculation is, at the present stage of our knowledge, an uncalled for and undesirable appendage to the systems of chemical formula notations.

Furthermore, it is not paradoxical for pragmatic chemists to advocate and employ metaphysical speculation, which within its proper limits is the soul of science, for, philosophically speaking, many of us believe that all science is metaphysics and necessarily so because, first, there is no such thing as a final explanation in any science; and, second, all science is established on empiricism only by and through sensation experience. Sensation experience is impossible without consciousness, and consciousness has never been defined by any philosopher, psychologist, or behaviorist.

Our present status with respect to our knowledge of the ultimate constitution of matter, that is of the atom, and the related nature of valency, was described in some lines written nearly two hundred and forty years ago, as follows:

Man doth with dangerous curiosity,
These unfathomed wonders try;
With fancied rules and arbitrary laws,
Matter and motion he restrains;
And studied lines and fictious circles draws;
Then with imagined sovereignty,
Lord of his new hypothesis he reigns.
He reigns; how long? 'till some usurper rise;
And he, too, mighty thoughtful, mighty wise,
Studies new lines, and other circles feigns.

Accordingly, while other circles are being feigned, the pragmatic chemist is disposed to advocate some such simple system of polar-

ity valence notation, as indicated in this paper, to correlate structural formulas and chemical behavior; and thus he prefers to avoid the imposition upon his structural formulas of unnecessary and irrelevant complications which, at the present status of his knowledge of the constitution of the atom, are more likely to confuse than to elucidate.

REFERENCES

- (1) STEWART, Recent Advances in Organic Chemistry, Longmans, Green and Co., (1927), Vol. II, p. 354
- (2) CRANSTON, J. Soc. Chem Ind, **47**, 208 (1928).
- (3) LEWIS, Valence and the Structure of Atoms and Molecules, Chemical Publishing Co., (1923)
- (4) FRY, The Electronic Conception of Valence and the Constitution of Benzene, Longmans, Green and Co, (1921).
- (5) LOWRY, Trans. Faraday Soc., **19**, 488 (1923).
- (6) NOYES, Trans Faraday Soc, **19**, 522 (1923).
- (7) KERMAK AND ROBINSON, J. Chem. Soc, **121**, 437 (1922).
- (8) MAIN SMITH, Chem. News, **135**, 81 (1927).
- (9) ABEGG AND BODLANDER, Z anorg. Chem., **20**, 453 (1899).
ABEGG, *ibid*, **39**, 330 (1904).
- (10) FRY, loc cit., Bibliographical Review of Published Applications, pp 272-285.
- (11) For later conceptions of W A Noyes, see his "The Relation between Shared Electrons and Valence; Principal and Contra Valences"—Sonderdruck ans. Z physik. Chem, Cohen-Festband, Akademische Verlags-gesellschaft m. b H in Leipzig, 1927.
- (12) LEWIS, loc cit., p 83.
- (13) TANASESCU, Bull soc. chim., **41**, 1468 (1927).
- (14) FRY, Z physik. Chem., **76**, 385 (1911); J Am. Chem. Soc., **37**, 855 (1915).
- (15) FRY, J. Am. Chem. Soc, **37**, 864 (1915).
- (16) LEWIS, loc cit, p. 85.
- (17) FRY, Z. physik. Chem., **76**, 387 (1911).
- (18) KENDALL, Inorganic Chemistry, Century Co. (1926), p. 535.
- (19) FALK, Private Communication, December, 1927.
- (20) J. J. THOMSON, The Electron in Chemistry, J. B. Lippincott Co. (1923).

THE THEORY OF PARTIAL POLARITY OF THE ETHYLENE BOND AND THE EXISTENCE OF ELECTRO-ISOMERISM

M. S. KHARASCH* AND F. R. DARKIS

Department of Chemistry, University of Maryland, College Park

INTRODUCTION

Many rules have been proposed regarding the addition of unsymmetrical reagents to compounds containing double bonds. A careful search of the literature, however, reveals numerous exceptions to any of the rules thus far formulated. In the present paper the writers propose a theory for the addition reactions of compounds of the ethylene type, which in their estimation is in agreement with the facts available in the literature. In the case of 2-pentene, where the evidence seemed to point against the theory proposed, the writers have been able to demonstrate the existence of two forms of the 2-pentene which add the halogen acid in a different ratio than that recorded in the literature.

GENERAL CONSIDERATION OF PREVIOUS THEORIES

a. Electronic theories proposed

With the advent of the electronic conception of valence, a number of theories appeared which attempted to explain and amplify the reactions of unsaturated derivatives as given in the rules of Markownikow, (1) Saytzeff (2), and Ipatiew¹ and also to account for the facts observed by other investigators. Only the outstanding theories will be considered here.

* Present address, Kent Chemical Laboratory, The University of Chicago, Illinois

¹ Ipatiew [Ber. **36**, 1988 (1903); Chem. Zentr. **70**, II, 17 (1899)] had a broad understanding of the subject for he believed that the position taken by the halogen, when a halogen acid is added to a double bond, is determined by the nature of the radicals linked to the carbon atoms of the double bond, the nature of the solvent, and the temperature.

The two electronic theories which have gained vogue are those of Lewis (3) and Cuy (4). Two attempts have been made to extend the Lewis theory by making specific assumptions, one by Carothers and the other by Lucas and his collaborators.

b. Discussion of theory of alternating polarity

The Cuy theory of alternating polarity has attracted much attention and has, because of its weakness, even been used as an argument against the application of the electronic concept to organic compounds. A careful perusal of Cuy's articles disclosed the weaknesses of the argument that there is an alternation in the charges on carbon atoms. The whole theory was shown, in the able paper of Lucas and Jameson (5), to be built upon erroneous assumptions. The strongest argument that Cuy believes he possesses, namely the alternation in melting points of homologous series, is undoubtedly due to differences in the crystal lattice structure. Thus, Verkade, Hartman and Coops (6) show definitely that the phenomenon of oscillation of heats of combustion of dicarboxylic acids of the oxalic acid series, observed by Stohmann, Kleber, Langbein and Offenbauer (7), holds up to C_{13} , the variation amounting on the average to 3 kgm. cal. per mole. (i.e., in one series the addition of CH_2 amounts to 158 kgm. cal., while in the other series it is only 155 kgm. cal.). The oscillation, however, disappears completely when the liquid esters of the same series are considered (8). Thermally the liquid esters constitute a single homologous series. Verkade, Coops and Hartman also point out that oscillation has been observed with fatty acids in the case of the following properties:

1. Melting point
2. Heats of crystallization (9)
3. Heats of combustion of crystalline acids (10)
4. Molecular volumes of crystalline acids (11)

but no oscillation has ever been observed with

1. Boiling points
2. Molecular refraction of the liquid acids (12)
3. Heats of combustion of liquid acids (10)
4. Molecular volumes of the liquid acids (11)

Hence the even and odd terms of the homologous series may have different crystal structures.² This view is rather an extension of the views of Garner and Randall (9), based on their work and also that of Muller and Shearer (13) on the heats of crystallization of fatty acids. There is thus no justification for this bit of evidence adduced by Cuy. The other three lines of argument, namely, (a) the addition of halogen acids to ethylene derivatives, (b) addition of halogen acids to hydrocarbons of the allene series, (c) rearrangement of alkyl bromides, have been shown by Lucas and Jameson to be without any justification.

These arguments against the theory of alternate polarity are reiterated at some length here for the hypothesis has taken root (14) and considerable experimental work was even done to disprove the hypothesis. It is also quoted quite often as an hypothesis of merit. We believe, however, in view of all that has been stated before, that the facts upon which this hypothesis is built are in themselves, when critically examined, the best argument against the theory.

The second electronic theory of the double bond was presented by Lewis (3) and elaborated upon by Carothers (15) and by Lucas (5) and his collaborators. However, before discussing these theories it would be well to state our own hypothesis so that we may show in what ways we depart from the views of these authors.

HYPOTHESIS OF THE DOUBLE BOND BETWEEN CARBON ATOMS FROM THE STANDPOINT OF THE THEORY OF PARTIAL POLARITY

a. Basis of theoretical concepts

The general basis for this concept of the double bond is discussed in papers by Kharasch and Sher (16) and Kharasch and Marker (17). In this first paper an attempt was made to correlate the electronic structure of organic substances with the total

² Cf. however, the paper by Pauly (*Z. anorg. allgem. chem.* **119**, 271 (1921) and also Garner and Randall (*loc. cit.*) The discussion of this paper is held in abeyance for Prof. Verkade promises to discuss it in a future publication. However, a critical review of these papers by the reader will convince him that the conclusions of Verkade, Coops and Hartman are in no way invalidated.

amount of heat they furnish upon combustion. The reference positions, to which the positions occupied by the electrons in all types of organic molecules are compared, are those occupied by the eight valence electrons around the carbon nuclei in methane and in carbon dioxide respectively. Furthermore, it is assumed that the displacement of electrons from the methane arrangement to that of carbon dioxide occurs in stages. This postulate is, of course, a corollary to the Bohr conception that around each atomic center there exists a particular set of energy levels. The postulate concerns itself only with those energy levels that lie outside the normal valence energy levels and the displacement of the electrons from the valence energy levels into outside energy levels. Hence reference to the eight electrons of methane is to their location in the normal valence energy level.³

The above postulate also implies that two atoms can share a pair of electrons in any of their respective energy levels and that the kinetic and potential energy of the system will depend upon the electronic structure of the molecule. To make clearer the meaning implied by this last statement, let the arcs in figure 1 represent some of the possible energy levels which electron pairs may occupy in carbon atoms *A* and *B*, respectively.

It is readily seen that two carbon atoms can hold a pair of electrons in various manners which entail varying degrees of polarity. The position of these shared electrons with respect to their individual valence orbits will depend upon the affinity of the two atoms for electrons and the nature of the other atoms attached to *A* and *B*. Thus, if the atom *A*, due to the nature of radicals attached to it has a greater affinity for electrons than does atom *B*, we may assume that the two electrons will be in the first

³ The adoption of this terminology does not commit us to any specific mechanical theory of the structure of the atom. If there existed a detailed and commonly acceptable model of the carbon atom, it would be possible to translate our theory into terms of that model. It would be still more convenient, however, for all practical purposes to employ a more schematic representation. Therefore, in directing the attention to figure 1, we particularly desire to emphasize the fact that no portrayal of actual electronic paths is attempted. The figure indicates only a sufficient number of energy levels to illustrate conveniently the concept we desire to advance.

energy level of *A* and the fourth energy level of *B* as indicated by diamonds in the figure. Other positions are also possible, for instance, the first energy level of *A* and the third energy level of *B*, as indicated by circles in the figure. The two atoms may thus share a pair of electrons in such a way as to give rise to various degrees of polarity. It is also quite possible that, if due to the nature of radicals attached to it, *A* has a slight affinity for electrons, it would tend to hold the electrons in some outer energy level, and if *B* has the same radicals attached to it, they would

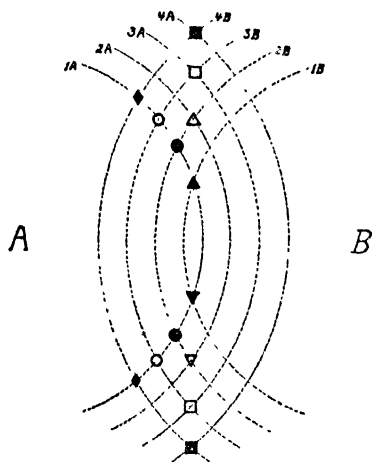


FIG 1 SCHEMATIC REPRESENTATION OF THE BOND BETWEEN TWO CARBON ATOMS AND SOME POSSIBLE VARIATIONS THEREOF

both tend to share the electrons in some outer energy level, as denoted by 4A and 4B in figure 1. This would undoubtedly contribute to the instability of the molecule.

b. Types of bonds in organic molecules and their heat value

We have thus, as far as the energy of the organic molecules are concerned, to consider that the possibilities shown in figure 2 exist in the sharing of electrons.

The lines in figure 2 represent the distances of the electrons forming the bond from the two carbon nuclei, A and B. We

assume also that the pair of valence electrons held together by two carbon atoms in saturated hydrocarbon i.e. type (1) is intermediate between their respective normal energy levels and the level which the electrons occupy in carbon dioxide. The net energy which they supply upon combustion is thus the same as if we had a pair of electrons in the valence shell of one carbon atom of the methane type. This rather plausible assumption is well justified by the agreement between calculated and observed values.

Accordingly if the expression for the heat of combustion of compounds of the type (1) is given by the expression $Q = 26.05 N$,

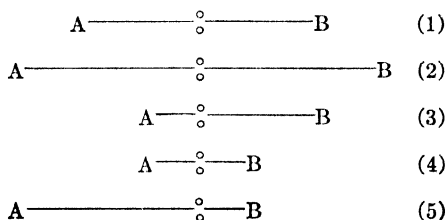


FIG. 2. TYPES OF BONDS IN ORGANIC MOLECULES

(1) Type of bond characteristic of aliphatic hydrocarbon, with ethane as a representative compound; (2) a bond in which both groups joined are weakly electronegative as in the case of the two carboxyls forming oxalic acid, or two triphenylmethyl nuclei; (3) a bond between a moderately electronegative radical (such as methyl) and a strongly electronegative radical like phenyl or naphthyl; (4) a bond between two highly electronegative radicals, such as phenyl or naphthyl; (5) a bond between a highly electronegative radical (such as phenyl or naphthyl) and a very weakly electronegative radical (such as carboxyl). Type 5 may be considered as a special case of type 3.

where N denotes the number of electrons, then if 2, 3, 4, 5 differ from (1) only in the arrangement of the one pair of electrons, the expressions for the heats of combustion of these compounds should be:

1. $Q = 26.05 N$
2. $Q = 26.05 N + a$
3. $Q = 26.05 N - b$
4. $Q = 26.05 N - c$, where c is larger than b
5. $Q = 26.05 N + d$

In case of compounds of type (1), the agreement within experimental error of the calculated heats of combustion of some thirty-

one saturated hydrocarbons with the experimentally determined values, indicates the plausibility and at least the approximate correctness of the hypothesis. This is merely a rough approximation to the truth, but until more refined experimental methods of attack are available the hypothesis covers the ground adequately.

This attack on the problem suggests immediately that in case of sharing of electrons as represented in type (2), the molecule contains a larger amount of energy, namely that necessary to pull the electrons out from the "normal" energy levels of atoms A and B into the outer energy levels. Upon combustion this molecule should furnish a larger amount of heat than calculated upon the assumption applicable to type (1). Unfortunately, the data at hand do not allow us to draw far-reaching conclusions. Thus, Schmidlin (18) determined the heat of combustion of a molecule which we should consider to have a bond of type (2) between the two carbon atoms, namely triphenyl methyl or hexaphenyl ethane. Our calculations, assuming the C-C bond to be similar to (1), give 2370.4 kgm. cal. for the heat of combustion, while Schmidlin obtained 2381.0 kgm. cal. The difference between the two values, namely 10.6 kgm. cal.⁴ could be ascribed to the potential energy accumulated in the molecule when the valence electron was moved out from the "normal" energy level into a position corresponding to bond (2). It is noticeable, however that the difference amounts to only 0.5 per cent, which could well be within the limits of experimental error of Schmidlin's determinations. The evidence is not sufficient to be admissible as proof of the validity of our argument. The purely chemical evidence, however, is much better and is in perfect agreement with our hypothesis.

The bond corresponding to type (3) and (4) is encountered when two strongly electronegative radicals share a pair of valence electrons. Due to the great attraction of each one of those radicals for electrons, they share these electrons more in the inner shells than do those which are not so strongly electronegative. The total energy which such molecules furnish upon

⁴ This difference is really even larger because the calculated value refers to the liquid state, while Schmidlin's value is that of the solid compound.

combustion would, therefore, be smaller, for it would require more energy to displace the electrons into the outermost energy levels. The radicals which would begin to exhibit that tendency are all those that appear above the methyl radical in our table of electronegativity of radicals (17).

This is the basis of the footnote on page 639 in the paper by Kharasch and Sher:

"It is most significant that for most benzene hydrocarbons the calculated value is somewhat higher than the experimentally determined value. This is due to a diminution of the energy of the system due to the benzene nucleus. Once this is taken into account in the calculation of the heat of combustion the agreement is then even more striking. It is omitted here in order to avoid confusion and will be discussed in detail in a later publication."

This, however, would imply a sharing of electrons in different energy levels for every different case, i.e. a formula which would express the sharing of electrons in different energy levels between a methyl and a phenyl radical would not express that for an ethyl and a phenyl radical. No doubt this notion is correct, but the increment is so slight as to be utterly incomparable to the normal experimental error. In bonds of type (3) and (4) *b* and *c* have the following numerical value: *b* = 3.5 kgm. cal., *c* = 6.5 kgm. cal. The value for *c* is larger than *b* because a bond of type (4) corresponds to a pair of valence electrons held in common by two strongly electronegative radicals.⁵

The heat of combustion of benzene hydrocarbons of the type (3) then becomes $Q = 26.05 \times N - 3.5A$, where *A* corresponds to a number of other carbon atoms which are directly linked to the benzene nucleus, and which lie below the phenyl radical in the table of electronegativity of radicals. Thus, *A* is 1 in toluene, 2 in xylene, 3 in mesitylene, etc. The calculated values agree remarkably well with the experimentally determined values, at all times well within the limits of experimental error. Similarly,

⁵ Similar considerations would undoubtedly apply to other types of compounds, which are made up of strongly electronegative radicals: phenols, aromatic ethers, etc

the heat of combustion of benzene hydrocarbons, corresponding to bond (4) is given by the expression $Q = 26.05 \times N - 6.5 B$, where B corresponds to the number of bonds between two electronegative radicals. Thus B is 1 in diphenyl, and 2 in p-diphenyl phenylene.

c. Heats of combustion of compounds containing double bonds and deductions therefrom

When the heat of combustion of a substituted ethylene derivative, say hexylene, is calculated on the same basis as was used for the heat of combustion of a saturated hydrocarbon,⁶ we obtain a value of $Q = 26.05 \times 36 = 937.8$ kgm. cal. per mole, while the experimentally determined value is 952.6 kgm. cal. Such a large difference between the calculated and the determined value is

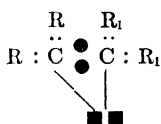


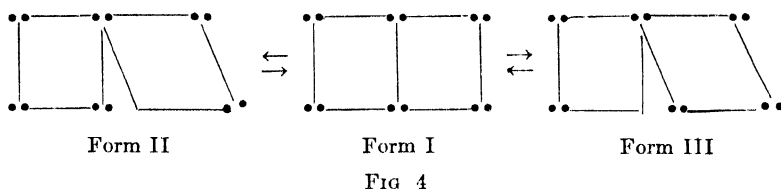
FIG. 3

significant in view of the fact that the agreement for the saturated hydrocarbons was within two tenths of one per cent. Furthermore, all unsaturated derivatives of the ethylene type give more heat upon combustion than that calculated by assuming that all the electrons are shared in a manner similar to that of the saturated hydrocarbons. The difference is almost constant within the limits of experimental error and amounts to about 13 kgm. cal. per mole.

Upon the basis of the view of molecular energy content developed, this difference is ascribed to the fact that the second pair of electrons held between the two carbon atoms is displaced from the normal valence energy levels of one or both of the carbon

⁶ The general formula for the heat of combustion with oxygen of a saturated hydrocarbon is given by the expression $Q = 26.05 N$, where Q is the molecular heat of combustion and N the number of electrons in the molecule, while 26.05 is the heat liberated by the inter-displacement of an electron from the methane "normal" energy level to that of the energy level of carbon dioxide type.

atoms. Furthermore, from the chemical evidence, which we shall adduce later, we conclude that there is a definite polarity at a double bond. Our total picture of the double bond in light of these facts may be formulated as in figure 3, where the circles denote the pair of valence electrons of unsaturated compounds which are of no consequence as far as addition reactions of the double bond are concerned. They may be bonds of the type 1, 2, 3 or even 4 depending upon the nature of the radicals R and R_1 . Their relative position would be of consequence only in the thermal behavior of the substance. The other pair of electrons, marked with squares, is the important one and the figure is intended to show that this pair of electrons is decidedly displaced from the carbon atom holding the R radicals so as to be in the



outer energy levels of that atom, and is not so far displaced from the normal valence orbit of the carbon atom holding the radicals R_1 . Of course, whether these electrons are in energy level 1, 2, 3 or 4 would depend upon the nature of the radicals R and R_1 .⁷

This is our general picture which, for the sake of clarity, is purposely devoid of such refinements which atomic structure information would allow us to add. However, to transform this into a useful hypothesis, we must add the following postulates:

d. Postulates of theory of partial polarity

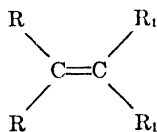
Postulate I. There is a definite polarity at the double bond, so that unsaturated derivatives of the ethylene type have a definite

⁷ It is evident from this concept that while we are employing the factor 13 kgm. cal. for the displaced electrons in ethylene compounds the latter value need not necessarily be the same for all types of unsaturated derivatives. Thus, it may be 9 kgm. cal. for some types and 16 kgm. cal. for others. The actual displacements cannot be determined within certain limits solely from consideration of heats of combustion, for the experimental errors involved are within these limits.

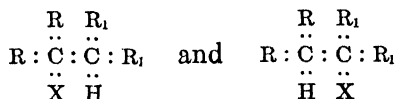
electrical polarization in at least one part of the molecule. We are rather reluctant to suppose that unsaturated derivatives, in which the radicals attached to the carbon atoms of the double bond are different, possess a structure as indicated in figure 4 and used by Carothers.

Forms II and III are assumed by Carothers to be capable of only momentary existence so that the concentration of the active forms is always small; the dissociations are assumed to be negligible so that there is equilibrium between the active forms and the inactive. We believe, and will prove later that there is no real evidence for the assumption that carbon compounds assume such an inactive form but that the directive addition reactions of unsaturated derivatives indicate that we have either Form II or III.

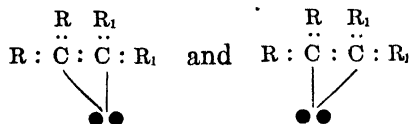
Postulate II. In the case where the radicals attached to the carbon atoms of the double bond are about the same in electronegativity (see Postulate VI) it is possible to produce either Form II or Form III. These two forms are electro-isomers. At no time do we assume, however, that these two forms are in equilibrium, through Form I. This postulate implies that if by the addition of a molecule of HX to an ethylene derivative



we obtain two products



and that they arose through the presence in the solution of two electro-isomers:



respectively. It is implied also, that one of the electro-isomers is more stable and that while it is possible to transform one electro-isomer into another, the two are not in equilibrium⁸ for the transformation of one into the other would be accompanied by a change in the energy of the system.

Postulate III. The second pair of electrons constituting the double bond is always held in outer energy levels of one of the carbon atoms and some closer inner energy levels of the other, the particular energy level depending upon the nature of the radicals attached to the double bond. The extent of this displacement from the arrangement of electrons in the methane molecule, we believe to be responsible for the larger energy content of ethylene derivatives. Furthermore, while this displacement in the case of the ethylene derivatives cited is about 13 kgm. cal. per mole, that may merely be due to the fact that the experimental error of determination of heats of combustion may be too large for the determination of such small differences, which may exist if the second pair of electrons is held in the second instead of the first energy level of figure 1, or even in the third, etc. At any rate, this concept implies that this second pair of electrons may be considerably displaced, but still shared in outer energy levels from one carbon atom, and displaced to some extent (as represented for instance by an amount of energy equivalent to 13 kgm. cal. which amount of energy is required to displace these electrons from their "methane" energy level arrangement) from the other "normal" valence energy level of the other. However, the extent of this latter displacement may vary considerably, depending upon the radicals attached to the carbon atoms, so that we may have a number of gradations of double bonds, a fact which will naturally manifest itself in the ease with which various double bonds will absorb reagents.

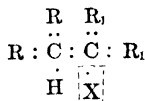
Postulate IV. The relative position of the second pair of valence electrons of the double bond depends upon the nature of radical attached to it in such a way that they are always on the carbon atom

⁸ This assumption is made for the sake of simplicity and also because there are no data which compel us to think differently. This assumption could, however, be verified if we had some data on the rate of addition of reagents to these electromers.

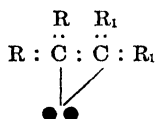
opposite to that carrying the most electronegative radicals. We were led to this assumption from the facts described in the paper of Kharasch and Marker, namely that the substitution of an electronegative phenyl radical for a hydrogen atom in the methyl group decreased the electronegativity of the radical, i.e., the fact that the methyl radical is more electronegative than the benzyl radical. Secondly a study of the addition reactions of simple ethylene derivatives containing decidedly electronegative radicals led us also to the postulate stated above. In speaking of relative electronegativity of radicals we make use of the table of electronegativity of organic radicals given in the paper of Kharasch and Marker and some general principles in regard to electronegativity of radicals which we shall develop later.

Postulate V. Alkyl and aryl radicals are more electronegative than a hydrogen atom. [This is not, strictly speaking, correct, for while it applies to simple radicals, it is conceivable that there are radicals less electronegative than hydrogen. Thus indications are that the radical CH_2Br or any other radical CH_2X where X is a strongly electronegative radical may be almost equal in electronegativity to the hydrogen atom. However, in the absence of a table of electronegativities including the hydrogen atom, we shall allow this to stand. Elaborations as to types of radicals which are less electronegative than hydrogen are given later in the paper.

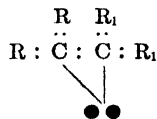
Postulate VI. In the formation of an unsaturated derivative from an alkyl halide it is assumed that the halogen is removed, taking with it the pair of valence electrons and the hydrogen drops off to maintain the electrical neutrality of the molecule, thus:



This gives Structure I for the unsaturated molecule,



Structure I



Structure II

which may or may not be the stable structure for this unsaturated derivative (see Postulate III). The following three alternatives are possible.

1. If the radical R_1 is decidedly more electronegative than the radical R the form given above (Structure I) would represent a stable structure.
2. If the radical R is decidedly more electronegative than the radical R_1 the following structure would probably be the structure of the compound, under the conditions that unsaturated derivatives are ordinarily prepared, namely, Structure II.

This of course represents a rearrangement of a pair of valence electrons induced by the strain in the molecule, and would occur only if, as stated, the radical R is very much more electronegative than the radical R_1 .

3. If the radical R is only very slightly more electronegative than the radical R_1 the compound of Structure I may exist, but when the molecule is subjected to the action of heat, or light, it will tend to rearrange to the most stable form, Structure II. It is among this class of compounds that one should presumably look for electro-isomerism, and where the isolation of electromers should be the easiest.

Postulate VII. In case the radicals attached to the carbon atoms of the double bond are all different the same principles apply as were developed for compounds in which two radicals of the same type were assumed to be attached to a carbon atom.

COMPARISON OF THE PARTIAL POLARITY HYPOTHESIS WITH THOSE OF LEWIS, CAROTHERS AND LUCAS

a. Carothers' hypothesis

Carothers' able paper (15) is an attempt to interpret some facts of organic chemistry by means of the Lewis-Langmuir concepts involving the octet theory. It is impossible, however, by using his premises to predict the addition reactions at the double bond.

b. Lucas' extension of the Lewis hypothesis

Lucas and his collaborators (19) follow closely in all the essentials the Lewis theory and attempt to adduce experimental evidence in favor of that hypothesis. Therefore, in discussing the work of these authors we will at the same time get a glimpse at the validity of the Lewis hypothesis.

As mentioned above, we fully agree with Lucas' criticism of the Cuy hypothesis, but we believe that the facts advanced by Lucas are not entirely valid, which would also invalidate the hypothesis advanced by Lucas and collaborators. These investigators reason on the basis of electron displacement, that if a molecule of 2-pentene is treated with HBr, the reaction should yield 3-bromopentene, since the methyl group exerts a stronger pull on electrons than ethyl group. The basis of this argument according to these authors is found in the fact that the methyl radical must exert a stronger attraction for the electrons than the ethyl radical for the ionization constant K_a for acetic acid is larger than K_a for propionic acid. Lewis also states "hydrogen is more negative than a methyl radical, but more positive than a phenyl radical. Methyl alcohol is a weaker acid than water, phenol is a stronger acid."

c. Validity of the concept of electron displacement

1. Ionization constants of acids. 2. Addition of unsymmetrical reagents to compounds containing double bonds. If rigidly applied the above argument leads nowhere. Thus formic acid should then be a weaker acid than benzoic acid while in reality it is a much stronger acid, the ionization constants being K_a formic 2.14×10^{-4} and K_a benzoic 6.6×10^{-5} . However, while this is a serious objection and would invalidate this method of determining electronegativity of organic radicals, yet we wish to raise a few other objections to the method, for we admit that by a superficial examination one may be tempted to draw such a conclusion. Unfortunately, the data at hand do not allow one to draw far-reaching conclusions. Thus, an examination of the data of the ionization constants of the aliphatic acids shows

sufficient variations to invalidate any conclusion drawn from it. The data merely indicate that ionization is molecular function. This latter property we believe to be shown more strikingly by a consideration of tables 1 and 2. Thus, in table 1 we note that there is a definite effect of the halogen atom in increasing the ionization constant of the acid, and that the effect

TABLE 1
Ionization constants, K_a for halogen acids

	Cl	Br	I
α -halogen propionic acid .	1.6×10^{-3}	1.08×10^{-3}	9×10^{-4}
β -halogen propionic acid	8.6×10^{-5}	9.8×10^{-5}	
α -halogen butyric acid	1.5×10^{-3}	1.06×10^{-3}	
β -halogen butyric acid	9×10^{-5}		
γ -halogen butyric acid .	3×10^{-5}	2.62×10^{-5}	1.7×10^{-5}
δ -halogen valeric acid . . .	1.91×10^{-5}	1.91×10^{-5}	
α -chloro crotonic acid .	7.2×10^{-4}		
β -chloro crotonic acid .	1.44×10^{-4}		
α -chloro isocrotonic acid	1.58×10^{-3}		
β -chloro isocrotonic acid	9.47×10^{-5}		

TABLE 2
Ionization constants K_a of some organic acids

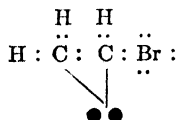
	Cl	Br	I
<i>Halogen benzoic acid</i>			
ortho	1.32×10^{-3}	1.45×10^{-3}	1.32×10^{-3}
para	9.3×10^{-5}	6.6×10^{-5}	1.36×10^{-4}
meta	1.55×10^{-4}	1.37×10^{-4}	1.63×10^{-4}
	ORTHO	META	PARA
Toluic acid.	1.25×10^{-4}	5.2×10^{-5}	4.5×10^{-5} *

* The values vary from 5.2 to 3.8

diminishes rapidly, and the drop is sudden after the α carbon atom, so that when the halogen is on β or γ carbon atoms the effect is relatively slight. These facts we admit to be in perfect agreement with the Lewis theory. However, a rigorous application of this principle to the aromatic acids, without assuming that it is the entire radical, R , rather than the halogen alone which

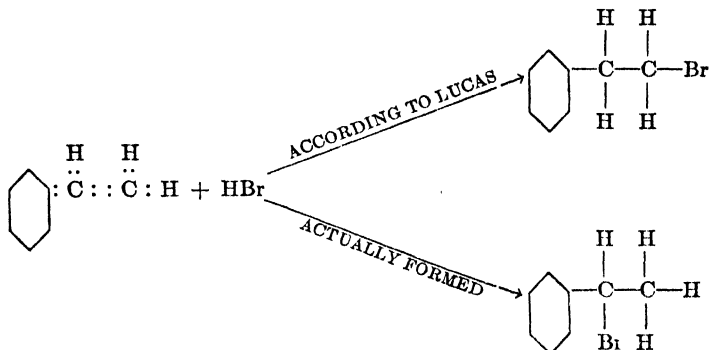
increases dissociation, would predict that in the case of the halogen benzoic acids the effect in increasing dissociation should be very slight since in the ortho acids the halogen is on the β -carbon atom. In reality the increase in strength of the acid corresponds to the increase observed in the aliphatic series when the halogen is on the α -carbon atom. Furthermore, the replacement of a hydrogen atom by a methyl group in the aliphatic series decreases the ionization of the acid, while when the methyl group replaces a hydrogen in the ortho position of benzoic acid, the strength of the acid is increased, while the same replacement of the hydrogen in the para and meta positions decreases the strength of the benzoic acid. The only conclusion that one may draw from these data is that ionization is a molecular effect, and that in our imperfect knowledge of the effect of the solvent in causing ionization and even the significance of our measurement, no reliance should be placed on this particular method of reasoning and that it is not permissible to draw conclusions in regard to the relative electronegativity of organic radicals from these premises.

Furthermore, it can be shown by a consideration of a number of organic addition reactions that this principle is not applicable. Thus, it is well known, that when vinyl bromide is treated with HBr, ethylidene bromide is formed. We repeated that reaction, and while we found the reaction to proceed slowly, yet the product formed was 99 per cent ethylidene bromide. Now, vinyl bromide according to Lewis and Lucas and collaborators should have this structure,



since bromine is more electronegative than a hydrogen. This compound should then add HBr to give ethylene dibromide—a conclusion in complete disagreement with the experimental fact. Furthermore, according to Lewis' concept since phenyl is more electronegative than hydrogen, styrene should add HBr to give

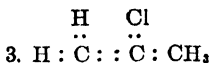
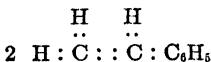
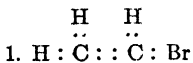
phenyl ethyl bromide, again in complete disagreement with the experimental fact, for the product formed is proven to be bromo ethyl benzene



Similarly, in the case of β -chloro propylene the theory of Lewis as elaborated by Lucas would predict that this product should add HI to give $\text{CH}_2\text{I}-\text{CHCl}-\text{CH}_3$ while the product recorded in the literature is $\text{CH}_3-\text{CClH}-\text{CH}_3$.

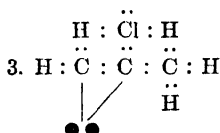
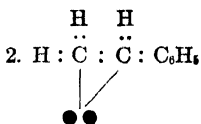
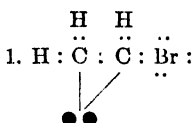
Perhaps these facts in themselves are sufficiently illustrative that the fundamental concept of electron displacement as developed by Lewis and extended by Lucas and collaborators is not valid and that it is impossible to set up a theory on the basis of these concepts which will adequately cover the data and help us systematize the known facts.

However, the examples cited fit very readily into the theory of partial polarity. Thus, writing down the examples which disagree with the Lewis theory,



and applying our postulates we are able to assign at once the position of the second pair of valence electrons and thus predict

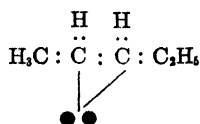
the course of the addition at the double bond. Thus, in examples 1, 2 and 3, each of the radicals attached to the double bond is more electronegative than the hydrogen atom, then according to Postulate IV, the second pair of valence electrons would be found on the carbon atom opposite the one carrying the most electronegative radical. The following formulas would, therefore, represent the polarity at the double bond and it may be readily noticed that the predicted addition reactions of these molecules are in conformity with the experimental data.



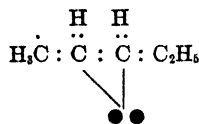
ELECTRO-ISOMERISM OF ETHYLENE COMPOUNDS

a. The existence of electromers of 2-pentene

There is one fact, however, which is advanced by Lucas and Moyse which is in disagreement with our theory, namely the addition of HBr to 2-pentene to give 22 per cent of 2-bromopentane and 78 per cent of 3-bromopentane. According to these authors the molecule of 2-pentene should have an electronic structure as represented by Form 1,

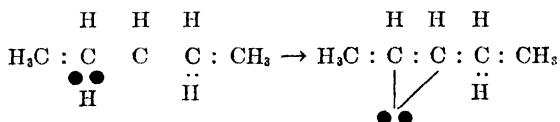


Form I



Form II

for a methyl radical is more electronegative than an ethyl radical and will attract the valence electrons toward itself. According to our theory, however, since methyl is more electronegative than ethyl, Form II should represent the position of the second pair of electrons (see Postulate IV). The experiment performed by Lucas and Moyse indicated, however, that 78 per cent of 3-bromopentane was formed and only 22 per cent of 2-bromopentane, more in agreement with Formula I than II. We repeated Lucas and Moyse's work and found their results to be entirely correct. A careful analysis, however, of Lucas and Moyse's method of preparing the 2-pentene revealed the fact that since the 2-pentene is prepared from 3-bromopentane the following might represent the course of the reaction:



The bromide ion is removed by the alcoholic potassium hydroxide, the hydrogen ion drops off to maintain the electrical neutrality of the molecule and unsaturated derivative of Form I would be produced. Also, since the methyl radical is only slightly more electronegative than the ethyl radical, the rearrangement to Form II, the most stable condition would only be a slow one, and that might account for the addition of HBr as observed by Lucas and Moyse, and confirmed by us. The possibility then suggested itself that, by applying a "strain" to the system, the transition might be expedited. From that it was reasoned that perhaps heat might transform the 2-pentene of Form I into Form II. Accordingly, the 2-pentene was prepared exercising the most scrupulous care, and the work of Lucas and Moyse checked repeatedly, using a Zeiss dipping refractometer (which reduced the error of observation from 8 to 3 per cent) and the conditions studied which might have had an effect on the values obtained by Lucas and Moyse. Table 3 indicates, for instance, that the percentage of HBr in the acetic acid mixture is of small consequence, and that once the procedure employed by Lucas and

Moyse is followed faithfully entirely reproducible results are obtained.

There was no question, therefore, about the validity of Lucas and Moyse's observation and, furthermore, it indicated that we were dealing with the same compound. However, when the xylene solution containing the pentene was heated to 100° for eight to sixteen hours, and the material then worked up as before different results were obtained as illustrated in table 4. This

TABLE 3
Checks on Lucas' and Moyse's work

NUMBER	REFRACTIVE INDEX	2 BROM COMPOUND	3 BROM COMPOUND	ACETIC ACID	HBr
		<i>per cent</i>	<i>per cent</i>	<i>grams</i>	<i>grams</i>
1	1 44371	22	78	60	43
2	1 44371	22	78	90	65
3	1 44371	22	78	84	80
*	1 44366	23	77	40	39

* Pentene subjected to five distillations before bromide was made.

TABLE 4
Pentene heated to 100°

NUMBER	INST READING	REFRACTIVE INDEX	2 BROM COMPOUND	3 BROM COMPOUND	TIME
			<i>per cent</i>	<i>per cent</i>	<i>hours</i>
1	79.1	1 44346	32	68	16
2	79.0	1 44343	32	68	16
3	79.1	1 44346	32	68	8
4	78.8	1 44337	34	66	16

difference is of course, beyond our experimental error, and furthermore, in many cases the two pentenes (heated and unheated) were recovered from the xylene solutions, in which they were kept, and worked up side by side, and invariably the 2-pentene which had been heated gave us a larger quantity of 2-bromo derivative than the unheated pentene.

Our next task was to ascertain the effect of temperature and light. The tubes containing the xylene solution of 2-pentene

were immersed in a glycerin bath and the latter exposed to light from a Mazda lamp. Table 5 indicates that temperatures of 50–60° had very little effect.⁹

Further investigation in regard to the existence of the two electromers of 2-pentene consisted in an endeavor to change the

TABLE 5
Pentene subjected to various temperatures in presence of light

NUMBER	INST READ- ING	REFRACTIVE INDEX	2 BROM COM- POUND	3 BROM COM- POUND	TIME HEATED	TEMPERA- TURE	SOLVENT	ADDED SUBSTANCE
			<i>per cent</i>	<i>per cent</i>	<i>hours</i>			
1	80 2	1.44377	22	78	8	50–60°	Xylene	Bromine*
2	79 3	1.44351	29	71	14 5	145–160°	Xylene	
3	79 1	1.44346	32	68	8	90–100°	Xylene	
4	79 3	1.44351	29	71	8	80–90°	Xylene	
5	79 3	1.44351	29	71	8	80–90°	None	

* A trace of bromine added to the pentene before heating had evidently no effect on the proportion of the two electromers formed.

TABLE 6
Pentene subjected to ultra-violet light

NUMBER	INST READING	REFRACTIVE INDEX	2 BROM COMPOUND	3 BROM COMPOUND	TIME EXPOSED	LENGTH OF TIME AFTER EXPOSURE UNTIL BROMIDE WAS MADE
			<i>per cent</i>	<i>per cent</i>	<i>hours</i>	
1	79 2	1.44348	30	70	6	Two days
2	78 5	1.44329	38	62	8	
3	78 5	1.44329	38	62	8	Two to four hours
4	78 1	1.44318	42	58	16	Two to four hours
5	77 4	1.44298	49	51	29	Two to four hours
6	78 4	1.44323	40	70	29	Three days
7	78 5	1.44327	38	62	29	Two days

proportion of the two forms (as judged by the different amounts of the two bromo derivatives formed when HBr is added to the mixture) by the use of ultra-violet light. The result of those

⁹ Whether the lower yield at 145°–160° is significant is difficult to decide. The difference is well within the limit of the error of the experiment. Yet, the yields of the 2-Bromo derivative were consistently lower at higher temperatures. This point is under investigation at the present time.

experiments was a large change in the proportion of the two electromers¹⁰ (table 6).

In examining table 6 it is important to bear in mind that the technique of HBr addition was not varied in any way from that employed when we were repeating the work of Lucas and Moyse. The only difference between the two experiments was that in one case the 2-pentene was exposed to ultra-violet light and some hours later HBr added to it under exactly the same condition as called for by Lucas and Moyse, the outcome of the addition, however, being entirely different. We attribute this difference to the fact that under the action of ultra-violet light the proportion of the two electromers was changed. The physical side seems also to lend some support to this view. Thus, in one set of experiments a large quantity of 2-pentene was put into a quartz tube and the gross absorption spectra determined. The 2-pentene was then exposed to the action of ultra-violet light and at stated intervals of eight, sixteen and twenty-nine hours, the absorption spectra determined again and at the same time samples were withdrawn and treated with hydrogen bromide. The absorption spectra work was only qualitative in nature, but a more quantitative study is under way. The absorption spectra work was carried out for us by Dr. B. H. Carrol of the Bureau of Standards, who has most graciously furnished us with the following report:

"The material said to be β -amylene was submitted with the request that the absorption spectrum be examined before and after exposure to ultra-violet radiation for varying lengths of time, as it was hoped that there would be a shift in absorption corresponding to the change in the relative amounts of the electromers in the sample. The sample was too volatile and too limited in amount to permit transfer to and from a cell, and the exposures covered a period of five days. It was, therefore, possible only to make approximate determinations of the

¹⁰ We believe this to be the first instance where the effect of ultra-violet light on a substance containing a double bond has been proved to consist merely in the shift of the relative position of electrons with respect to the two carbon atoms. This notion has some far-reaching possibilities particularly with regard to the effect of ultra-violet light on vitamin D (Ergosterol). The relationship of our findings to vitamin D potency is under investigation by one of us.

absorption limit after each period of exposure. The tube containing the sample was used as a condenser to focus the radiation from a mercury arc on the slit of a Fery quartz spectrograph; for each determination three exposures of varying length were made with the tube in place, and three with the bare arc. The development of the films was comparable in all cases. The results given are based merely on visual observation of the lines.

"The method used permits only the statement that a shift in the limit of absorption to the longer wave lengths has been produced by exposure. This indicates an irreversible change in the composition of the sample; since, applying Le Chatelier's principle, an equilibrium should be shifted in the direction of decreasing absorption when the shift is due to absorption of radiant energy."

TABLE 7

TIME OF EXPOSURE	LONGEST WAVE LENGTH COMPLETELY ABSORBED	SHORTEST WAVE LENGTH TRANSMITTED	LONGEST WAVE LENGTH AT WHICH ABSORPTION CAN BE DETECTED	2 BROM COMPOUND	3 BROM COMPOUND
<i>hours</i>	<i>Å</i>	<i>Å</i>	<i>Å</i>	<i>per cent</i>	<i>per cent</i>
0	2753	2804	3028	22	78
8	2753	2804	3139	38	62
16	2857	2894	3663	42	58
29	2926	2968	3663	49	51

The last two columns of table 7 have been added by us in order to indicate that chemically, as judged by proportion of the two bromides formed when HBr is added to the substance, the 2-pentene has been undergoing a change when exposed to the action of ultra-violet light. We might add that we are inclined to interpret the data on addition reactions where two products are formed by the addition of a reagent to a substance containing a double bond, as arising only from the existence of two distinct electromers of that substance.

b. Does heat or ultra-violet light effect a change in the position of the double bond?

In order to prove conclusively that Electromer II produced from I is not 1-pentene, i.e., that heating the substance in an

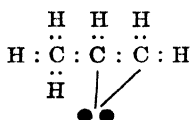
inert solvent did not shift the double bond to position-1, the two products were treated separately with bromine and the dibromides thus formed were compared. As shown in the experimental part (to be published elsewhere) there was not the slightest difference between the compounds thus formed—the dibromides agreeing perfectly in their indices of refraction, boiling point and density. We may safely assume then that heating did not shift the position of the double bond, but merely shifted the valence electrons in the sense previously discussed.

c. Are the electromers cis-trans isomers?

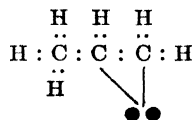
It is readily noticed that 2-pentene may exist in two forms, *cis* and *trans*. It may, therefore, be argued that the two electromers are *cis-trans* isomers. The data at hand do not allow one to draw such a far reaching conclusion. For that would imply that *cis-trans* isomers would add an unsymmetrical reagent in a different manner. The fact that the bromine addition products are the same would argue against such a conclusion. However, bromine is a poor reagent for such a purpose, and the oxidation of these electromers with permanganate—a reagent which is less likely to cause shifts from *cis* to *trans* is now being studied.

d. Pertinent remarks regarding isolation of other electromers

As previously discussed, the isolation of electromers of the type indicated should be possible whenever the radicals attached to the carbon atoms of the double bond are very close in electronegativity. However, if there is a great deal of difference in their electronegativity, then, even though we may attempt the preparation of electromers by eliminating the groups in the proper order, the rearrangement into the stable electromer will take place very rapidly. Thus, we attempted to prepare the two theoretically possible electromers of propylene by elimination of HBr from normal and isopropyl bromides respectively.



Electromer I



Electromer II

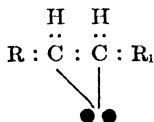
However, the addition of HBr to the propylenes prepared by these methods resulted under the experimental conditions, approximating those that we employ for the two pentenes, in the formation of isopropyl bromide exclusively, a fact which tends to show that although Electromer I of propylene was undoubtedly formed, it rearranged itself to II very quickly, due to the large difference in electronegativity between methyl group and the hydrogen atom.¹¹

Similarly, in the case of styrene one would not expect to isolate the two possible electromers under conditions ordinarily employed in the laboratory, for the difference in electronegativity between the two radicals phenyl and hydrogen, is so large that the unstable electromer would rearrange very rapidly into the other form.

EXTENSION OF THE THEORY OF PARTIAL POLARITY TO ALL TYPES OF ETHYLENE DERIVATIVES

a. Predictions regarding addition of unsymmetrical reagents to ethylene compounds and criteria thereof

The directed addition reactions at the double bond may be predicted by making use of the postulates given, once the relative electronegativities of the radicals have been definitely established, particularly in cases where the effect of two radicals in displacing the valence electrons of the double bond is in the same direction. Thus, if in the molecule:



the radical R is strongly electronegative, while R_1 is very weakly electronegative, the position of the valence electrons of the double bond would be as indicated, and only one product would result from the addition of halogen acid. However, if both radicals R and R_1 are electronegative then the position of the

¹¹ In our analysis of the propyl bromides formed we employed the index of refraction. A preliminary examination of the index of refraction of known mixtures of propyl and isopropyl bromides indicated that the index of refraction was a linear function of composition.

second pair of valence electrons would depend upon the relative positions of the radicals R and R_1 in the table of electronegativity, i.e. they would be closer to the nucleus of the carbon atom carrying the least electronegative radical. However, even though that form would predominate the possibility is that the other form would also exist, and two products should be formed, when HBr is added to such a molecule.¹²

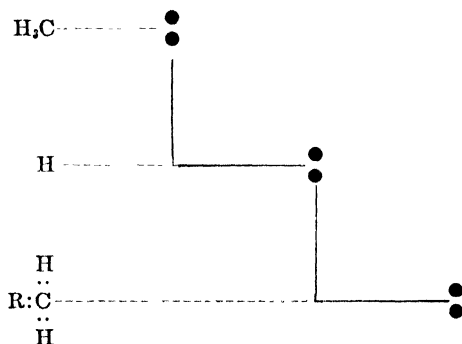
The main difficulty of predicting all types of addition reactions is as previously mentioned, due to our ignorance of the relative electronegativity of organic radicals. However, with two simplifying assumptions it becomes possible to reconcile our addition reactions at the double bond with the facts actually known. The first assumption is that any alkyl or aryl radical is more electronegative than hydrogen. Secondly, that if a carbon atom has strongly electronegative radicals attached to it that group may become equal or lower in electronegativity than the hydrogen atom. It is among such compounds that one may find a large number of examples of electro-isomerism, and it is among such compounds that we find that experimental conditions will determine to a large extent the relative amounts of the two compounds which are always formed.

The first assumption is self evident—it means that such groups as phenyl, methyl, ethyl or naphthyl are more electronegative than the hydrogen atom, although it is quite conceivable that an overlapping may occur somewhere, for instance in the case of long-chain aliphatic radicals. In the second assumption we merely imply that since an electronegative radical attached to a methyl carbon atom decreases enormously the electronegative character of the group, that it would tend to put us on the other side of the electronegativity of the hydrogen atom.

¹² Once this view is adopted it becomes possible to compare the relative electronegativity of such radicals as phenyl and bromine. The literature records that the product of the reaction of bromo-styrene with HBr gives styrene dibromide. Our own experience with this reaction led to the same result. This would seem to indicate that the phenyl radical is more electronegative than a bromine atom. Work is under way at the present time to determine the relative electronegativity of a phenyl radical with a chlorine and fluorine atom, by studying the addition reactions of halogen acid to the corresponding styrene halogen compounds.

Thus, while the phenyl radical, a bromine, or chlorine atom, or an hydroxyl group or oxygen atom are individually more electronegative than a hydrogen atom, when they become attached to a methyl carbon atom as for example in the radical benzyl or bromo methyl or chloro methyl they make those groups either less electronegative than hydrogen, or almost equal in electronegativity to the hydrogen atom. The general idea of this concept is conveyed in table 8.

TABLE 8
Possible electronegativity of radicals with respect to hydrogen

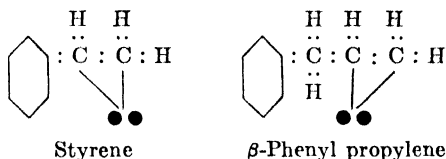


Where *R* is a strongly electronegative radical

In addition reactions at the double bond, therefore, such groups as benzyl or bromo or chloro methyl, carbonyl, or carboxyl when attached to a carbon atom of the double bond should have an effect opposite to the one that the phenyl radical or bromine atom have when attached to the same carbon atom of the double bond. The prediction would be that in vinyl bromide the bromine atom forces the second pair of electrons on to the opposite carbon atom. However, in allyl bromide, the position of the pair of valence electrons will depend upon the relative electronegativity of the two radicals, hydrogen atom and bromo methyl radicals. These radicals are quite close together in the electronegativity series. According to our postulate, however, since the hydrogen atom is slightly more electronegative, the second pair of

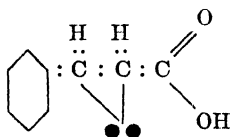
electrons of the double bond will be closer to the carbon atom carrying the bromo methyl group than the carbon atom carrying the hydrogen atom. The system, however, will be a labile one, and capable of change by the application of heat or other forms of energy—such as light.

Our experiments confirmed these suppositions, for when vinyl bromide and allyl bromide were treated with HBr they gave only ethylidene bromide and 1,3 dibromo propane (when the addition was carried out in light). It was to be expected also from these considerations that while styrene adds HBr to give 1-bromo ethyl benzene, the molecule β -phenyl propylene containing a group attached to a carbon of the double bond which is presumably less electronegative than hydrogen should have this structure:



and therefore add hydrogen bromide to give γ -phenyl normal propyl bromide. Experiment confirmed this prediction. However, since the bromide was not described in the literature, the product of the reaction was shaken with silver oxide. The alcohol obtained in this manner was γ -phenyl normal propyl alcohol.

Similarly, since the carbon atom of the carboxyl group has strongly electronegative atoms and groups attached to it, with the consequent displacement of all the electrons of the carbon atom into outer energy levels (Postulate I), the whole group is an exceedingly weak electronegative radical, weaker than the hydrogen atom. This would imply that if the carboxyl group and a bromine atom or a phenyl or methyl group are attached to the two carbon atoms of the double bond respectively, their effect would be in the same direction and only one product should result. In the case of cinnamic acid the electronic structure at the double bond would be, upon this hypothesis, the following:



the pair of valence electrons being completely over in the direction of the carbon atom next to the carboxyl, although they would still be shared in some outer energy levels of the carbon atom carrying the phenyl radical.

On the other hand, if the strongly electronegative radical and the carboxyl group are attached to the same carbon atom of the double bond their effect would be in the opposite direction, and we might expect two products. It is also among such compounds that one may look for the phenomenon of electro-isomerism.

b. Theoretical explanation of ease of addition of reagents by compounds containing double bonds

The halogens add to the double bond in the order of chlorine, bromine, iodine, while the reverse order holds for the halogen acids. This is easily accounted for on the basis of the general concept of the double bond represented in figure 2. For it is at once evident that while we may consider the first step in the absorption reaction of a halogen acid to consist in the addition of the hydrogen ion to the displaced electrons, before the addition is complete, the electrons from the halogen must be pulled into the inner energy levels of the carbon atom which is deficient in electrons. Naturally it requires larger quantities of energy to accomplish this in the case of the chloride ion, than in the case of the bromide and iodide ions. It follows also that the more electronegative the radicals R the farther out will be the electrons which the carbon atom of the double bond will be capable of holding—the smaller would be the amount of energy necessary to pull the electrons from the halogen, and the greater should be the ease of addition. We believe this to be in harmony with the described experimental facts. Thus it is recorded that while ethylene does not add HCl , chloroethylene adds it readily. Other illustrations are available in the literature. However, it must be

remembered at all times that ease of addition implies that the valence electrons of that carbon are more readily satisfied in outer energy levels and with a strongly electronegative halogen attached to it, the molecule should also decompose more readily. It is quite possible that in a good many cases where we state that the substance does not add a reagent, it may be due to the fact that while addition does take place, the decomposition of the molecule under the conditions of the experiment is exceedingly rapid, and that under another set of conditions the molecule may actually be isolated.¹³

SUMMARY

A comparison is made between the various electronic hypotheses which have been advanced to interpret the directed addition reactions of ethylene derivatives. It is shown that the hypothesis of partial polarity of the double bond¹⁴ covers the existing data more satisfactorily than any of the other theories. The application of this hypothesis to thirty-five ethylene addition reactions recorded in the literature gives results in conformity with the experimental facts.

REFERENCES

- (1) MARKOWNIKOW, *Ann* **153**, 256 (1870)
- (2) SAYTZEFF, *Ann* **179**, 297 (1875).
- (3) G N LEWIS, *Valence and the Structure of Atoms and Molecules*, A. C. S. monograph, The Chemical Catalog Co, New York, 1923
- (4) CUY, *J Am Chem. Soc* **42**, 503 (1920)
- (5) LUCAS AND JAMESON, *J Am Chem. Soc.* **46**, 2475 (1924)

¹³ Note, for instance that at low temperatures triphenyl ethylene adds bromine, but as soon as the temperature rises HBr is evolved. The net result is the replacement of a hydrogen atom by a bromine atom, but it is significant and in accord with our theory that the substitution takes place through addition. The instability of the molecule is also in agreement with our predictions

¹⁴ From the standpoint of the hypothesis developed in the paper one can not agree with the arguments and deductions of Gilman and Peterson (*J. Am Chem. Soc* **48**, 423 (1926) with regard to the non-existence of a polar bond in ethylene derivatives. The non-addition of a reagent (in this case the Grignard reagent) has no bearing to polarity in the molecule—but may be due to unfavorable energy change if a combination of that type took place. Furthermore the experiment as designed did not take into account the true polar ionization of the Grignard reagent.

- (6) VERKADE, HARTMAN AND COOPS, *Rec trav Chim* **45**, 373 (1926).
- (7) STOHMANN, KLEBER, LANGBEIN AND OFFENHAUER, *J. prakt Chem.* (2) **49**, 117 (1894).
- (8) VERKADE, COOPS AND HARTMAN, *Rec trav Chim* **45**, 586 (1926).
- (9) GARNER AND RANDALL, *J Chem Soc* **125**, 881 (1924).
- (10) VERKADE AND COOPS, unpublished work (private comm).
- (11) GARNER AND RYDER, *J Chem Soc* **127**, 720 (1925)
- (12) EYKMAN, *Rec trav. Chim* **12**, 157 (1893)
- (13) MULLER AND SHEARER, *J Chem Soc* **123**, 2043, 3156 (1923)
- (14) HENRICH, *Theories of Organic Chemistry*, transl by Johnson and Hahn
John Wiley and Sons, Inc , New York, 1922
- (15) CAROTHERS, *J. Am Chem Soc.*, **46**, 2226 (1924)
- (16) KHARASCH AND SHER, *J Phys. Chem* **29**, 625 (1925)
- (17) KHARASCH AND MARKER, *J. Am Chem Soc* **48**, 3130 (1926)
- (18) SCHMIDLIN, *Ann chim phys* (8), **1**, 195 (1906)
- (19) LUCAS AND JAMESON, *J Am Chem Soc* **36**, 2475 (1924); LUCAS, SIMSON AND
CARTER, *ibid* , **47**, 1462 (1925); LUCAS AND MOYSE, *ibid* , **47**, 1459 (1925).

THE BROADER ASPECTS OF VALENCE: ITS APPLICATION TO COLLOID COMPOUNDS

VICTOR COFMAN

Experimental Station, E. I. du Pont de Nemours & Company, Wilmington, Delaware

The classical idea of valence was developed to explain the properties and behavior of "pure" chemical compounds.¹ Even in that field it does not at present completely account for the facts. As Sidgwick puts it: "The modern distinction of different types of valency shows that the combining power of an atom may be due to more than one cause and may arise in more than one way, and it is not possible to express all the variations of this power as numerical values of a single constant" (1).

CHEMICAL COMPOUNDS AND COLLOID COMPOUNDS

True chemical compounds are usually crystalline or crystallizable substances, of constant composition, with discontinuous properties (entirely different from those of their constituents). They are brittle when solid and usually not very sensitive to light.

On the other hand colloids are typically gel-like, have curved forms, variable composition, and properties changing continuously with outside conditions. They are generally elastic and

¹ The valence concept was originally introduced to account for certain whole number relations in the combining ratios of atoms and groups. A. W. Williamson's "monad" and "diad" groups, Frankland's "affinities" (1852), Gerhardt's "atomicity" (1853), Odling's "replaceable values" and Kekulé's "saturation capacity" (1857), laid the foundations of the doctrine of valence. Since that time numerous types of valence have been described: Polar, electro- or ionizable; non-polar, neutral or non-ionizable; normal or positive, contra- or negative; principal and subsidiary, primary and secondary, auxiliary, coordinate; absolute or maximum, active, actual, free or effective, passive, dormant, crypto- or unsaturated; electrical-double, dative or co-valence; conjugate, semi-, partial, virtual partial, latent, residual, potential; unsymmetrical, forced, absolute polar, homopolar, heteropolar; and no doubt many others

One can devise a scheme—as in table 1²—showing the transition from definite chemical compounds, through dissociable compounds, hydrates and alloys, hydrous oxides and oxysalts, elastic gels and adsorption compounds, solutions and dispersions—on to coarse physical mixtures. In many of the above systems it is difficult to state whether the forces that hold the atoms together are of the valence type or are due to some other kind of affinity.

For convenience' sake one may consider the current picture of the true chemical combination—namely, the transference or sharing of electrons. In sodium hydride the sodium is assumed to have lost an electron to hydrogen, the latter being therefore negatively charged:



This is chemical combination.

Now, consider a negatively charged gold particle containing adsorbed sodium ions (see fig. 1). Here, too, the sodium ions have a deficiency of electrons and the gold particle has an excess of electrons. But what we know of the freedom of movement of electrons in metals constrains us to admit that the excess electrons do not belong to any particular atoms of gold, but to the particle as a whole. Consequently one must say that there are cases in which the primary valence forces of a group of ions are satisfied by the electrical field of force of another group of ions, in non-equivalent proportion.

² Table 1 indicates that properties change gradually from one group to the next and no property belongs exclusively to one group (2). For instance, the change in energy accompanying formation—often supposed to be typical of chemical compounds,—is in reality common to all groups, but varies greatly in degree. Again, physical mixtures are characterized by ease of separation with simple methods involving gravitational or centrifugal forces; but gravitation will also cause partial separation of dispersions or even solutions, and it is possible that in some celestial bodies at high temperatures it may break up compounds owing to difference in the weight of their component atoms.

It is equally true that a substance of given composition may exist in different states. Chemical compounds, at high temperatures, may dissociate completely and become physical mixtures, while colloidal substances, under certain conditions, may assume crystalline form.

THE "LAW" OF CONSTANT PROPORTIONS

How does this conclusion affect the law of constant proportions? It merely shows it to be one more of those "laws of nature" which state that there are three feet to the yard. If, out of the infinite variety of substances in nature one chooses to call chemical compounds only those in which stoichiometric relations exist, one has made a useful definition. By saying, further, that substances can only combine chemically in definite proportions, one merely *restates* the definition of chemical compounds. This restatement has been given the form of a "law" and invested with possibilities of good and evil.

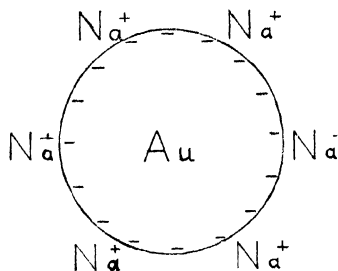


FIG. 1. COMBINATION (ADSORPTION OF IONS) IN NON-STOICHIOMETRIC PROPORTIONS

Consider its good effect first: Coming at a time when technical skill and knowledge of the properties of matter were rather limited, it was a distinct advantage to restrict the field of investigation of chemists to substances of constant composition. Such substances can generally be crystallized and are readily identifiable by definite properties. Even today with a greatly increased knowledge of nature and with the help of specialized apparatus, investigators find themselves baffled by many colloid problems. Had our predecessors attempted to study these difficult systems instead of the readily purified chemical compounds, they may well have become discouraged.

The law of constant proportions enabled them to build a truly magnificent edifice of chemical knowledge and to prepare and

describe hundreds of thousands of chemical species. However, in the enthusiasm of their great success, forgetting that they simply discovered in Nature only what they were looking for, chemists endowed the statement of constant proportion with universal validity; later, when they began to study less definite substances like proteins, soap, cellulose, etc., the ideas of constant proportion and discontinuous valence were so deeply entrenched, —and they still are,—that to some chemists it seems almost sacrilegious to discuss the limitations of these concepts (2). Yet a simple analogy will show how arbitrary such a view is.

Suppose that some superpower were interested to know how much a chemist drinks, in other words, the affinity of, say, American chemists for water, beer or wine. If this superpower were to follow the example set by chemists it would proceed somewhat along the following lines: It might begin with a very large number of chemists and start a race, rejecting those who reach the goal first and those who are left behind, retaining only those that run at about the same speed. This would correspond to fractional distillation, and would eliminate many of the stout chemists and leave a more homogeneous crowd. Next, this superpower might resort to crystallization by arranging the chemists into such patterns that it would necessarily place those of the same size together, and thus eliminate those that are too tall, or too short. After a few more "purifications" which would render the group as uniform as possible, the superpower would be ready to investigate the drinking capacity of the individuals; however, instead of doing so with each one separately this superbeing would calculate how many barrels the group, as a whole, needs. Under such circumstances the probability is that, no matter what group of chemists this superbeing began with, it would obtain the same *per capita* consumption of liquid; thus it could prove as a law of nature, that the combining affinity of chemists for water, or beer, or wine, has a constant value.

This analogy may seem fantastic, but proofs of the constancy of composition of many compounds (e.g., sodium caseinate or hydrated magnesium oxides) are, in my opinion, no less forced. The methods which are used in the so-called purification of these

substances, and the restrictions placed upon the conditions of the system are such as to make the law of constant composition necessarily true. No one would deny that the beautiful crystalline substances and the constant-boiling liquids that adorn the shelves of the chemical laboratories are examples of stoichiometric compounds and that their structure can be simply and plausibly explained by our present views of valence. There is no sufficient reason, however, to assume that theories which have been found useful for those specially prepared substances will be found convenient for the rest of nature. From the fact that telegraph poles are of standard sizes, it is not safe to generalize that all trees in the forest grow to the same height.

For the sake of uniformity and simplicity we should, of course, apply as far as possible, the same pictures and theories to all substances in nature. It should be apparent, nevertheless, that the differences between typically crystalloidal and typically colloidal substances are sufficiently great to warrant a new and more flexible viewpoint.

COMBINING RATIO (? VALENCE) AS A CONTINUOUS VARIABLE

From a purely experimental point of view, the valence or combining ratio of atoms, or groups, can be determined from the molar ratio of substances liberated at the electrodes on electrolysis. With chemical compounds this ratio is generally a simple whole number, but with colloidal compounds, like proteins and soaps, the ratio of acid to base liberated by the current varies with experimental conditions (pH, concentration, etc.). This abnormality is usually explained, as due to the presence of micells or complex ions (3). Independent of these views, however, it is possible to deal with this variable ratio—which, when constant, is called valence—as a continuous variable, depending on external conditions.

In the case of ions attached to a charged particle (fig. 1) the affinity—or valence, if one chooses to extend the meaning of that word—can also vary. The number of ions associated with the particle will depend on the charge of the particle, therefore on its electrical potential. If this increases above a certain

value—the critical potential of the adsorbed ion—then the latter will have its charge neutralized and will be liberated in the atomic state.

Another instance of adsorption, which is even more difficult to distinguish from chemical combination is afforded by a monomolecular film of oleic acid floating on water containing a little sodium hydroxide. It has been proven by Hartridge and Peters (4) that the amount of sodium hydroxide combined with the oleic acid varies continuously with the pH—and Rideal (5) has even calculated this so-called combination from Gibbs' law of surface adsorption.

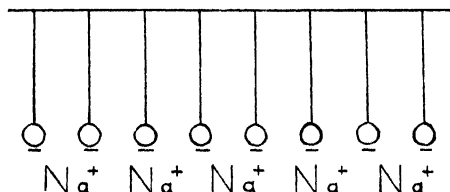


FIG. 2. ADSORPTION OF SODIUM IONS BY OLEIC ACID FILM FLOATING ON WEAK NaOH SOLUTION

The aptitude to combine in all ratios is not restricted to surface adsorption. Figure 3 shows how the composition of a soap gel or solution varies with pH. There is no reason whatever to suppose that the neutralization of oleic acid by sodium hydroxide is not uniform, in these instances. In other words, the composition of the gel and the affinity of sodium ions for oleate ions depends on the pH, which is an electrochemical potential.

The idea that combining capacity might be considered from a dynamic point of view, occurred to the writer originally from a different set of considerations: There are numerous physical laws which can be conveniently expressed in terms of an intensity factor, like pressure, and an extensive factor, like volume. These laws tell us how a change in the value of one of these factors, in one phase or system, affects the corresponding factor in a related system. For instance, the change in vapor pressure of a liquid

(dp) with change in total pressure on the liquid (dp') may be expressed thus

$$Vdp = V'dp'$$

where V and V' are the specific volumes of vapor and liquid respectively.

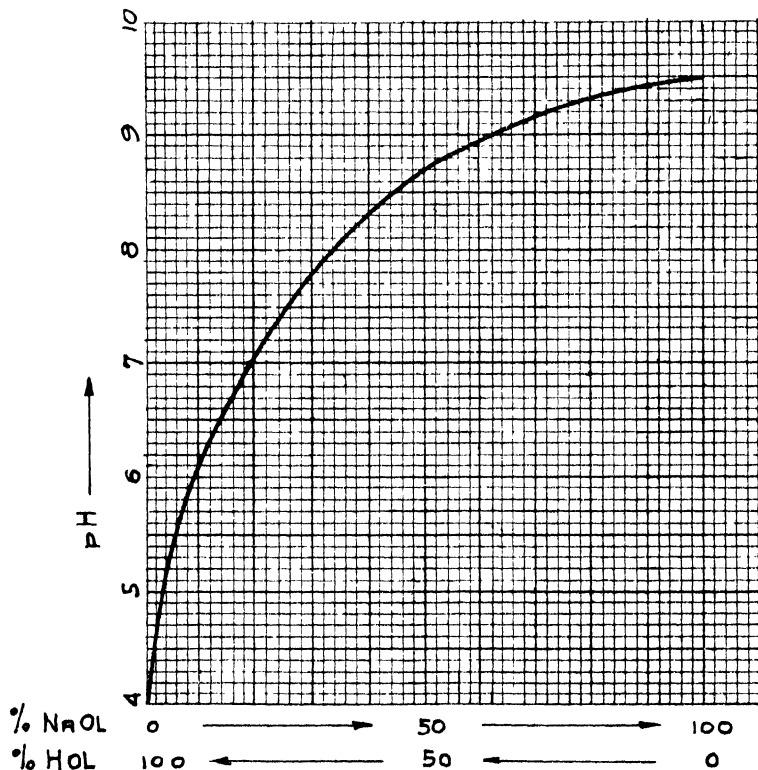


FIG. 3. RELATION BETWEEN pH OF 0.3 PER CENT SODIUM OLEATE SOLUTION AND PER CENT COMPOSITION

Note: pH determined by means of indicators

The relations between vapor pressure and osmotic pressure or swelling pressure, of osmotic pressure and surface pressure, can be expressed in an identical manner (6).

Similarly the relation between electromotive force and ionic pressure is given by

$$(Fn)dE = Vdp$$

where dE = change in pressure of electrons (e.m.f.).

dp = change in pressure of molecules or ions.

F = Faraday's constant.

n = valence; it corresponds to gas volume or space distribution of molecules and may, indeed, be considered a function of the *space distribution of electrons*.

Those who care to have a picture in terms of electrons may imagine that in the case of colloid compounds the electrons lost by the positively charged atom do not go to a single other atom but are shared, or revolve around a number of atoms or molecules. With this crude picture as a beginning one would be led to inquire about the orbits, the number, and the frequency of these electrons that move around several molecules. Investigations on particle size, extensibility and sensitivity to light of various wave length would naturally follow.

THE PRACTICAL IMPORTANCE OF A THEORY

Let us suppose that two young chemists are asked to carry out research work in an industry dealing with colloidal substances like rubber, soap, cement or cellulose. One of these young men has been taught, and believes, that atoms can only join in definite ratios. The other believes that they can combine in all proportions. How will these hypotheses influence their work?

We know what the classically trained chemist would do. If engaged in the investigation of soaps, he would begin to purify and separate various fatty acids and prepare their salts; he would succeed eventually in obtaining crystals of sodium oleate, dioleate, and tetra oleate. If working with gelatin, the stoichiometrist would try to obtain crystals and spend most of his time in proving that such compounds as sodium gelatinate and gelatin chloride, have constant ratios between their constituents (7). If

he were in the nitrocellulose industry, the man who believes in constant ratios would continue to discuss mono-, di-, and tri-nitrocellulose and attempt to prepare them pure, as others have done for the last half century.

Now, this point of view is entirely correct. By trying hard enough, and long enough, one may succeed in preparing, under special conditions, crystals of gelatin chloride (8) or nitrocellulose. Unfortunately, such crystals would not have the properties of the common substances which go by that name.

A soap manufacturer may have been in business a lifetime without ever having seen crystals of sodium oleate. He is not interested in that aspect of his material; therefore he does not appreciate the efforts of the classical chemist and trusts the production in the hands of technical men. On the other hand, it so happens that the man who takes the view that atoms can combine in all proportions will necessarily be led to investigate those properties of colloid systems which are of greatest practical value to biology and industry. He will inquire how the ratio between oleic acid and alkali affects the properties of the system, e.g., its pH, its power of absorbing water, its elasticity, etc.

In biology he would find out the relation between the amounts of the constituents and factors like pH and rH (oxidation-reduction potential), swelling pressure and other important physiological facts.

In dealing with cellulose and nitrocellulose the careful investigator would not simply determine, say, the copper value, and forget that the intensity of oxidation is just as important as the amount of oxidation. He might be led to inquire into the intensity of nitration of nitrocellulose rather than the relative proportions of mono-, di-, and tri-nitrocellulose. Also, as already pointed out, his attention would be forcibly drawn towards properties like polymerization, elasticity and sensitivity to light.

It has been retorted, with perfect truth, that there is nothing in the doctrine of valence or the law of constant proportions that will prevent a man from carrying out the same kind of experiments; that, in fact, much of the advance in our knowledge of colloids has been made by people who held strictly orthodox

stoichiometric views. This is true but not sufficient. If one man has learned Chinese and another French, there is nothing to prevent the one who knows Chinese from going to France or the one who speaks French from visiting China. But the immense probability is that, if they had leisure and means, the one who has read about China will cross the Pacific to see the Chinese wall while the other who has studied French history will cross the Atlantic to see the Cathedral at Rheims. There is no question here as to whether Chinese or French is the truer language, and similarly, the point is not whether valence and combining power are continuous or discontinuous. The essential fact is that these two views direct our attention to different problems.

TABLE 2

	BOUNDED BY PLANAR FACES	SOLID	SHOW DOUBLE RE- FRACTION	GIVE X-RAY DIFFRACTION IMAGE
True crystals	+	+	(+)	+
Fine grained rocks or crystalline powders		+	+	+
Liquid crystals			+	(+)
Stretched gelatin, rubber, etc			(+)	+

CRYSTAL FORMS AND COLLOID FORMS

As an example of the tendency to apply to colloid systems ideas derived from the study of chemical compounds—and in doing so to neglect the intrinsic properties of colloids—we may consider the use of the word “crystal” in connection with substances like rubber or cellulose. The term “crystal” originally defined a solid substance bounded by plane faces. Subsequent work with the microscope showed that most crystals are double-refracting and, more recently, Laue and Bragg have shown that all crystals give definite x-ray diffraction patterns. Mineralogists found it convenient, in the identification of minerals, to rely more upon optical behaviour than external shape as a criterion of crystallinity. Further, Lehmann (9) showed that certain liquids exhibit a property in common with crystals, namely double-refraction, and the term “liquid crystals” was used to show this similarity. Finally x-ray patterns were ob-



FIG. 4 COLLOID FORMS OBTAINED WITH ETHYL CELLULOSE

tained from rubber and some other colloids, frequently under tension, which showed that a certain similarity exists between colloids and crystals (10). The x-ray diffraction pattern can be obtained by the uniform arrangement of very small particles that are not necessarily bounded by plane faces. Thus the x-ray diffraction method is more inclusive than the former methods of judging crystallinity, and the use of limiting terms to describe the original conception of crystallinity becomes necessary. G. L. Clark (11) has recognized this by defining two states: "the *paracrystalline* state, involving the true, incipient and transition

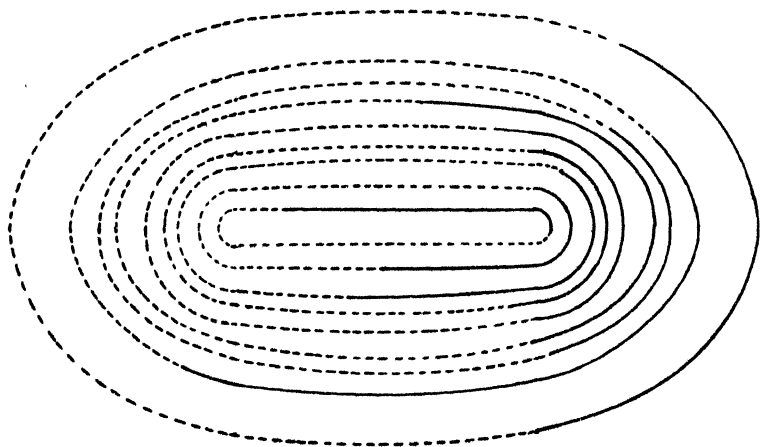


FIG. 5. REGULARITY OF COLLOID FORMS

processes which ultimately yield the true crystal with its permanent symmetrical disposition in direction and space" and "the *metacrystalline* state, a transient symmetrical arrangement of diffracting centers under special external conditions, which is not a transition towards a true crystal." All that the x-ray patterns indicate is a certain regularity in the arrangement of the constituent particles—and this is but one out of four important properties of a true crystal. It would be hardly correct to state that a book is crystalline because its leaves are arranged in definite order.

The chief objection to the use of the word "crystalline" in connection with colloid compounds like cellulose is that it directs our attention to straight lines and solid angles. In reality, a curved shape appears to be typical of colloids, just as the crystal is the normal form of a chemical compound. Figure 4 shows colloid forms obtained in films of ethyl cellulose.³ It will be seen that the curves can be fitted in a series of ellipses with common foci (fig. 5). This suggests that colloid forms follow as definite and simple laws as crystalline forms do, and reveals a valuable avenue of approach to the study of the shapes of bio-colloids such as leaves and flowers,⁴ a field which has been completely ignored by stoichiometrists. "Of all problems in physiology that of form is least approachable" states Warburg (12) and, we may add, "from the classical point of view."

CONCLUSION

The history of science supplies many examples of the limited usefulness of theories. A new theory is generally based upon a few simple hypotheses; it coordinates known facts and leads to new knowledge. As the number of facts to be explained by a given theory increases, it is often found that the original simple hypotheses no longer suffice. This is the position reached by valence theories with reference to the study of colloids.

The study of colloid behaviour is of relatively recent growth. Many of its most important aspects like the laws that regulate colloid forms and the action of light on colloids have hardly been touched upon. It is important that we approach this study with an open mind. If necessary, let us be ready to accept the viewpoint of variable affinity which broadens our concept of valence by removing an unnecessary restriction. It is probable, moreover, that we shall eventually be constrained to give up

³ These were obtained on pouring a 10 per cent dispersion of ethyl cellulose in butyl acetate on a glass plate, so as to form a layer 3 mm thick, and evaporating at 60°C

⁴ Examples of leaf-like forms taken up by inorganic salts (NH_4Cl , MgSO_4 , etc.) are illustrated in Kapacewski's book "L'état colloidal et l'Industrie," Paris 1927, p 282.

our cherished views of the electron as a definite particle, and picture it as a system of waves defined by a set of equations. It is to be expected that these new ideas will agree better with a continuously variable affinity than with a fixed valence.

REFERENCES

- (1) SIDGWICK: The Electronic Theory of Valence, Oxford (1927), p. 182.
- (2) W. KOPACZEWSKI: In Alexander's Colloid Chemistry, vol. 1, 547 (1926).
W. HERZ: Ahrens Sammlung, 19, 417 (1912).
- (3) J. W. MCBAIN AND R. C. BOWDEN. J Chem Soc, 123, 2417 (1923).
- (4) HARTRIDGE AND PETERS: Proc Roy. Soc (London) A 101, 348 (1922).
- (5) RIDEAL. Surface Chemistry, Cambridge (1926), p. 250.
- (6) COFMAN. Chem Reviews, 4, 26, 40 (1927).
- (7) A. L. FERGUSON: Coll Sympos Monograph, 5, 173 (1928).
J. LOEB: Proteins and the theory of colloidal behaviour, McGraw-Hill (1922).
- (8) C. H. BRADFORD: Biochem J, 14, 91 (1920), Trans Faraday Soc, 16, Append p. 56 (1920-21).
- (9) LEHMANN: Flüssige Krystalle, Leipzig (1904).
- (10) CLARK: Applied X-rays, N. Y., McGraw-Hill (1927), p. 174.
- (11) Ibid., p. 176.
- (12) O. WARBURG: J Cancer Research, 1, 143 (1925).

